Comparative Effectiveness Review

Number 7

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

This report is based on research conducted by the RTI International-University of North Carolina Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-02-0016). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of the Agency for Healthcare Research and Quality or of the U.S. Department of Health and Human Services.

Comparative Effectiveness Review

Number 7

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

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-02-0016

Prepared by:
RTI International-University of North Carolina Evidence-based Practice Center

Investigators
Gerald Gartlehner, M.D., M.P.H.
Richard A. Hansen, Ph.D.
Patricia Thieda, M.A.
Angela M. DeVeaugh-Geiss, M.S.
Bradley N Gaynes, M.D., M.P.H.
Erin E. Krebs, M.D.
Linda J. Lux, M.P.A.
Laura C. Morgan, M.A., M.P.H.
Janelle A. Shumate, M.D., M.P.H.
Loraine G. Monroe
Kathleen N. Lohr, Ph.D.

AHRQ Publication No. 07-EHC007-EF
January 2007

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

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

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the State Children's Health Insurance Program (SCHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting Comparative Effectiveness Reviews of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strengths and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see http://effectivehealthcare.ahrq.gov/reference/purpose.cfm.

AHRQ expects that Comparative Effectiveness Reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

Acknowledgments

This study was supported by Contract 290-02-0016 from the Agency for Healthcare Research and Quality, Task No. 7. We acknowledge the continuing support of Beth Collins-Sharp, Ph.D., R.N., and Lia Snyder, M.P.H., of AHRQ.

We extend our appreciation to the members of our Technical Expert Panel, listed below. All provided thoughtful advice and input during our research process.

The investigators deeply appreciate the considerable support, commitment, and contributions of the EPC team staff at RTI International and the University of North Carolina. We particularly acknowledge Timothy Carey, M.D., M.P.H., for insightful comments on the first draft and helpful suggestions throughout the project; Lynn Whitener, Dr.P.H., M.S.L.S., our EPC Librarian; and UNC supporting staff Stacey Williams and Jonathan Ward. We also express our gratitude to RTI staff Tammeka Swinson, Research Analyst, and both Jennifer Drolet, M.A., and Melissa Fisch, B.A., editors.

Technical Expert Panel

Susan G. Kornstein, M.D.
Professor of Psychiatry and Obstetrics and Gynecology
Virginia Commonwealth University
Richmond, Virginia

Anand Kumar, M.D.
Professor of Psychiatry and Director of Geriatric Ambulatory Care Programs
University of California
Los Angeles, California

John Williams, M.D.
Professor of Medicine and Psychiatry
Duke University
Durham, North Carolina

AHRQ Contacts

Beth A.Collins Sharp, Ph.D., R.N.
Director
Evidence-based Practice Center Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Rockville, MD

Lia Snyder, M.P.H.
Task Order Officer
Evidence-based Practice Center Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Rockville, MD

Contents

Executive Summary 1

Introduction 19
Background 19
Purpose of this Report 20
Scope and Key Questions 24

Methods 29
Topic Development 29
Literature Search 29
Study Selection 30
Data Extraction 30
Quality Assessment 31
Applicability Assessment 32
Rating Strength of a Body of Evidence 32
Data Synthesis 33

Results 35
Overview of All Key Questions 35
Key Question 1: Efficacy or Effectiveness in Treating Depressive Disorders and Symptoms 35
Major Depressive Disorder (MDD): Overview 36
Major Depressive Disorders: Key Points 44
Major Depressive Disorder: Detailed Analysis 48
Dysthymia: Overview 66
Dysthymia: Key Points 66
Dysthymia: Detailed Analysis 67
Subsyndromal Depressive Disorders: Overview 68
Subsyndromal Depressive Disorders: Key Points 68
Subsyndromal Depressive Disorders: Detailed Analysis 68
Key Question 2: Efficacy or Effectiveness for Maintaining Remission or for Treating Patients With Unresponsive or Recurrent Disease 69
Overview 69
Maintaining Remission: Key Points 71
Maintaining Remission: Detailed Analysis 71
Treating Treatment-Resistant Depression or Relapse or Recurrence: Key Points 78
Treating Treatment-Resistant Depression or Relapse or Recurrence: Detailed Analysis 78
Key Question 3: Efficacy or Effectiveness for Treating Symptoms Accompanying Depression 80
All Symptoms: Overview 80
Anxiety: Key Points 81
Anxiety: Detailed Analysis 84
Insomnia: Key Points 87
Insomnia: Detailed Analysis 88

Melancholia: Key Points 90
Melancholia: Detailed Analysis 91
Pain: Key Points 91
Pain: Detailed Analysis 93
Psychomotor Change: Key Points 94
Psychomotor Change: Detailed Analysis 94
Somatization: Key Points 95
Somatization: Detailed Analysis 95
Key Question 4: Comparative Harms and Adherence for Second-Generation Antidepressants 96
Overview 96
Adverse Events and Discontinuation Rates: Key Points 96
Adverse Events and Discontinuation Rates: Detailed Analysis 97
Severe Adverse Events: Key Points 105
Severe Adverse Events: Detailed Analysis 108
Adherence: Key Points 112
Adherence: Detailed Analysis 112
Key Question 5: Efficacy, Effectiveness, and Harms for Selected Populations 113
Overview 114
Age: Key Points 114
Age: Detailed Analysis 117
Sex: Key Points 122
Sex: Detailed Analysis 122
Race or Ethnicity: Key Points 123
Race or Ethnicity: Detailed Analysis 123
Comorbidities: Key Points 124
Comorbidities: Detailed Analysis 124

Discussion 129
General Conclusions 129
Results for Efficacy and Effectiveness in Major Depressive Disorders 136
Results for Maintaining Response or Remission 137
Results for Managing Treatment-Resistant or Recurrent Depression 137
Results for Treating Patients with Depression and Accompanying Symptoms 138
Results for Harms (Adverse Events) and Adherence 140
Results for Population Subgroups 140
Results for Dysthymia and Subsyndromal Depression 141
Future Research 141
Efficacy and Effectiveness 141
Prevention of Relapse and Recurrence 141
Management of Treatment-Resistant or Recurrent Depression 142
Accompanying Symptoms 142
Adverse Events 142
Addendum 143

References 145

Tables

1. Second-generation antidepressants approved for use in the United States 21
2. Usual dosing range and frequency of administration for adults 23
3. Criteria for effectiveness studies 25
4. Outcome measures and study eligibility criteria 26
5. Dosing classification based on lower and upper dosing range quartiles 32
6. Definitions of the grades of the overall quality of evidence 33
7. Abbreviations and full names of diagnostic scales and other instruments 37
8. Study characteristics, response and remission rates, and quality ratings of studies in adults with major depressive disorder 39
9. Possible comparisons of second-generation antidepressants involving SSRIs and number of included head-to-head trials 43
10. Possible comparisons of second-generation antidepressants involving SSNRIs, SNRIs, and other antidepressants and number of included head-to-head trials 44
11. Characteristics of trials comparing mirtazapine to SSRIs on onset of action (response rate) 46
12. Characteristics of trials comparing bupropion to SSRIs on sexual functioning and satisfaction 47
13. Characteristics and effect sizes of studies comparing citalopram with escitalopram 48
14. Interventions, numbers of patients, results, and quality ratings of studies in adults with dysthymia 66
15. Interventions, numbers of patients, results, and quality ratings of studies in adults with subsyndromal depressive disorders 68
16. Number of head-to-head comparisons and placebo-controlled studies for assessment of relapse and recurrence 70
17. Head-to-head studies of relapse prevention and recurrence prevention 72
18. Placebo-controlled studies of relapse prevention and recurrence prevention 74
19. Head-to-head studies of treatment-resistant and recurrent depression 79
20. Studies of adults with major depressive disorders and accompanying anxiety 81
21. Studies of adults with major depressive disorders and accompanying insomnia 87
22. Studies of adults with major depressive disorders and accompanying melancholia 90
23. Studies of adults with major depressive disorders and accompanying pain 92
24. Studies of adults with major depressive disorders and accompanying psychomotor change 94
25. Studies of adults with major depressive disorders and accompanying somatization 95
26. Studies assessing adverse events and discontinuation rates 98
27. Mean incidence of specific adverse events across comparative trials 99
28. Average rates of overall discontinuation, discontinuation because of adverse events, and discontinuation because of lack of efficacy 103
29. Studies assessing severe adverse events 106
30. Head-to-head trials reporting adherence to second-generation antidepressants 113
31. Studies of efficacy, effectiveness, and harms for patient subgroups 115
32. Summary of findings with strength of evidence 130

Figures

1. Phases of treatment for major depression 22
2. Results of literature search 36
3. Relative risk meta-analysis of MADRS response rates comparing citalopram with escitalopram 49
4. Effect size meta-analysis comparing citalopram with escitalopram on the MADRS 50
5. Relative risk meta-analysis of response rates comparing fluoxetine with paroxetine on the HAM-D 52
6. Effect size meta-analysis comparing fluoxetine with paroxetine on the HAM-D 52
7. Relative risk meta-analysis of response rates comparing fluoxetine with sertraline on the HAM-D 54
8. Effect size meta-analysis comparing fluoxetine with sertraline on the HAM-D 55
9. Relative risk meta-analysis of response rates comparing fluoxetine with venlafaxine on the HAM-D 58
10. Effect size meta-analysis comparing fluoxetine with venlafaxine on the HAM-D 58
11. Relative risks of response rates comparing SSRIs with SSRIs on the HAM-D 63
12. Relative risks of response rates comparing SSRIs, SNRIs, SSNRIs, and other second-generation antidepressants with other second-generation antidepressants on the HAM-D 64
13. Relative risks of response rates comparing SSRIs with SSNRIs and SSRIs with SNRIs on the HAM-D 65
14. Relative risk of nausea and vomiting of venlafaxine compared with SSRIs 101
15. Relative risks of overall discontinuation 104
16. Relative risk of discontinuation because of adverse events 104
17. Relative risk of discontinuation because of lack of efficacy 105

Appendixes

Appendix A. Peer Reviewers
Appendix B. Search Strategy
Appendix C. Excluded Studies
Appendix D. Evidence Tables
Appendix E. Characteristics of Studies with Poor Internal Validity
Appendix F. Placebo Studies Included in Meta-Regression
Appendix G. Placebo Studies Excluded from Meta-Regression
Appendix H. Meta-analyses of Discontinuation Rates
Appendix I. Publications Appearing Only as Abstracts
Appendix J. Acknowledgments

Executive Summary

Background

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

Pharmacotherapy dominates the medical management of depressive disorders and may include first-generation antidepressants (tricyclic antidepressants and monoamine oxidase inhibitors) and more recently developed second-generation antidepressants. These second-generation treatments include selective serotonin reuptake inhibitors (SSRIs) and serotonin and norepinephrine reuptake inhibitors (SNRIs). The mechanism of action of most of these agents is poorly understood. These drugs work, at least in part, through their effects on neurotransmitters such as serotonin, norepinephrine, or dopamine in the central nervous system.

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

This report summarizes the available evidence on the comparative efficacy, effectiveness, and harms of 12 second-generation antidepressants: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine in treating patients with MDD, dysthymia, and subsyndromal depression. It also evaluates the comparative efficacy and effectiveness for maintaining remission and for treating accompanying symptoms such as anxiety, insomnia, or neurovegetative symptoms.

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

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

1a. For adults with MDD, dysthymia, or subsyndromal depressive disorders, do commonly used medications for depression differ in efficacy or effectiveness in treating depressive symptoms?

1b. If a patient has responded to one agent in the past, is that agent better than current alternatives at treating depressive symptoms?

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

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

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

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

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

4. For adults with a depressive syndrome, do commonly used antidepressants differ in safety, adverse events, or adherence? Adverse effects of interest include but are not limited to nausea, diarrhea, headache, tremor, daytime sedation, decreased libido, failure to achieve orgasm, nervousness, insomnia, and more severe events including suicide.

5. How do the efficacy, effectiveness, or harms of treatment with antidepressants for a depressive syndrome differ for the following subpopulations:

4. Elderly or very elderly patients;

5. Other demographic groups (defined by age, ethnic or racial groups, and sex);

6. Patients with medical comorbidities (e.g., ischemic heart disease, cancer)?

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

Conclusions

Treatment of Major Depressive Disorder (KQ 1)

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

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

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

Seventy-two head-to-head comparisons (i.e., comparisons between medications conducted within trials) provided data on 35 of the potential comparisons between the 12 second-generation antidepressants addressed in this report. Five trials directly compared any non-SSRI second-generation antidepressant with any other non-SSRI second-generation antidepressant; of these, only one comparison was evaluated in more than one trial. Many efficacy trials were not powered to detect statistically or clinically significant differences, leading to inconclusive results.

Direct evidence from head-to-head trials was considered sufficient to conduct meta-analyses for four drug-drug comparisons. Differences in efficacy reflected in some of these meta-analyses are of modest magnitude and clinical implications remain to be determined.

- Citalopram vs. escitalopram (five studies; 1,545 patients): Patients on escitalopram had an additional treatment effect of a 1.25-point reduction (95-percent confidence interval [CI], 0.10-2.39) on the Montgomery-Asberg Depression Rating Scale (MADRS) compared with patients on citalopram. The relative risk (RR) of response was statistically significantly greater for escitalopram than for citalopram (RR: 1.14; 95-percent CI, 1.04-1.26). The number needed to treat (NNT) to gain one additional responder at week 8 with escitalopram was 14 (95-percent CI, 7-111). Both drugs are produced by the same manufacturer, which funded all available studies.

- Fluoxetine vs. paroxetine (seven studies; 950 patients): We did not find any statistically significant differences in effect sizes on the Hamilton Depression Rating Scale (HAM-D) or response rates between fluoxetine and paroxetine. Fluoxetine had an additional reduction of 0.55 (95-percent CI, –1.4-0.36; $P = 0.23$) points on HAM-D compared with paroxetine; paroxetine led to a higher rate of responders than fluoxetine (RR 1.09; 95-percent CI, 0.99-1.21).

- Fluoxetine vs. sertraline (four studies; 940 patients): Patients on sertraline had an additional, statistically nonsignificant treatment effect of a 0.75-point reduction (95-percent CI, –0.45-1.95) on the Hamilton Rating Scale for Depression (HAM-D) scale compared with patients on fluoxetine. The relative risk of response was statistically significantly greater for sertraline than for fluoxetine (RR: 1.11; 95-percent CI, 1.01-1.21). The NNT to gain one additional responder at 6 to 12 weeks with sertraline was 14 (95-percent CI, 8-22).

- Fluoxetine vs. venlafaxine (eight studies; 1,814 patients): Patients on venlafaxine had an additional, statistically nonsignificant treatment effect of a 1.31-point reduction (95-percent CI, 0.10-2.39) on the HAM-D scale compared with patients on fluoxetine. The relative risk of response was statistically significantly greater for venlafaxine than for fluoxetine (RR: 1.12; 95 percent CI, 1.01-1.24). The NNT to gain one additional responder at 6 to 12 weeks with venlafaxine was 12 (95-percent CI, 7-50). All studies were funded by the makers of venlafaxine.

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

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

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

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

Quality of life. Quality of life or functional capacity was infrequently assessed, usually as a secondary outcome. Eighteen studies (4,050 patients), mostly of fair quality, indicated no statistical differences in efficacy with respect to health-related quality of life. The strength of evidence is moderate.

Speed of response. Seven studies, all of fair quality and funded by the maker of mirtazapine, reported that mirtazapine had a significantly faster onset of action than citalopram, fluoxetine,

paroxetine, and sertraline. The NNT to yield one additional responder after 1 or 2 weeks of treatment is 7 (95-percent CI, 5-12); after 4 weeks of treatment, however, most response rates were similar. Again, this treatment effect was consistent across all studies, but whether this difference can be extrapolated to other second-generation antidepressants remains unclear. The strength of evidence is moderate.

Response to a second agent. The Sequenced Treatment Alternatives to Relieve Depression (STAR-D) trial is the only well-done study looking at the question of response to a second agent among those failing initial therapy. Results show that about one in four of the 727 people who participated in the switch became sym—bupropion sustained release (SR), sertraline, and venlafaxine extended release (XR).

Treatment of Dysthymia

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

One good-quality and four fair-quality placebo-controlled trials provide mixed evidence on the general efficacy and effectiveness of fluoxetine, paroxetine, and sertraline for the treatment of dysthymia. A fair-quality effectiveness study provides mixed evidence on the effectiveness of paroxetine compared with placebo. A subgroup of patients older than 60 years showed a significantly greater improvement than those on placebo; a subgroup of patients younger than 60 years did not show any difference in effectiveness between paroxetine and placebo. The strength of evidence is low.

Treatment of Subsyndromal Depression

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

Maintenance of Response or Remission (KQ 2a)

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

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

Efficacy and effectiveness. One head-to-head efficacy study and two effectiveness studies provide conflicting evidence on differences among second-generation antidepressants in treatment-resistant depression. The efficacy study (fair quality) suggests that venlafaxine is modestly more effective than paroxetine. A good-quality effectiveness study suggests that no substantial differences exist among bupropion SR, sertraline, and venlafaxine XR, but a fair-quality effectiveness study suggests that venlafaxine is modestly more effective than citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline. Given the conflicting results, the overall strength of the evidence is moderate.

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

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

Anxiety. Evidence from six head-to-head trials and one placebo-controlled trial (all fair quality) suggests that antidepressant medications do not differ substantially in antidepressive efficacy for patients with MDD and anxiety symptoms. The trials found no substantial differences in efficacy between fluoxetine, paroxetine, and sertraline; sertraline and bupropion; and sertraline and venlafaxine. One trial found statistically significant superiority of venlafaxine over fluoxetine. The strength of evidence is moderate.

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

Melancholia. Two head-to-head trials (both fair quality), one poor-quality head-to-head trial, and one fair-quality placebo-controlled study provide limited evidence on the comparative effects of medication for treating depression in patients with melancholia. In one, depression response rates for sertraline were superior to those for fluoxetine; in another, depression scores improved more for venlafaxine than for fluoxetine. The strength of evidence is low.

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

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

Somatization. We identified no relevant study.

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

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

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

Melancholia. We identified no relevant study.

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

Psychomotor changes. We identified no relevant study.

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

Differences in Harms (Adverse Events) (KQ 4)

We analyzed adverse events data from 72 head-to-head efficacy studies on 16,780 patients, along with data from 39 additional studies of both experimental and observational design**.** Only five RCTs were designed primarily to detect differences in adverse events. Methods of adverse events assessment in efficacy trials differed greatly. Few studies used objective scales. Determining whether assessment methods were unbiased and adequate was often difficult.

General tolerability.

Adverse events profiles. Constipation, diarrhea, dizziness, headache, insomnia, nausea, and somnolence were commonly and consistently reported adverse events. On average, 61 percent of patients in efficacy trials experienced at least one adverse event. Nausea and vomiting were found to be the most common reasons for discontinuation in efficacy studies. Overall, second-generation antidepressants have similar adverse events profiles, and the strength of evidence is high. However, some differences in the incidence of *specific* adverse events exist, as follows:

- Venlafaxine was associated with an approximately 10-percent (95-percent CI, 4-17 percent) higher incidence of nausea and vomiting than SSRIs as a class. In addition, pooled discontinuation rates because of adverse events in efficacy trials are statistically significantly higher for venlafaxine than for SSRIs (RR: 1.50; 95-percent CI, 1.21-1.84). The strength of evidence is high.

- In most studies, sertraline led to higher rates of diarrhea than comparator drugs (bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine). The incidence was 8-percent (95-percent CI, 3-11 percent) higher than with comparator drugs. Whether this finding can be extrapolated to comparisons of sertraline with other second-generation antidepressants remains unclear. The strength of evidence is moderate.

- Mirtazapine led to higher weight gains than comparator drugs (fluoxetine, paroxetine, venlafaxine, and trazodone). Mean weight gains compared to pretreatment ranged from 0.8 kg to 3.0 kg after 6 to 8 weeks of treatment. Paroxetine had higher weight gains than fluoxetine and sertraline. The strength of evidence is moderate.

- Trazodone was associated with an approximately 16-percent (3-percent less to 36 percent higher) higher incidence of somnolence than comparator drugs (bupropion, fluoxetine, mirtazapine, paroxetine, venlafaxine). Whether this finding can be extrapolated to comparisons of trazodone with other second-generation antidepressants remains unclear. The strength of evidence is moderate.

- Discontinuation syndromes (e.g., headache, dizziness, nausea) occurred in 0 to 86 percent of patients. Paroxetine and venlafaxine had the highest incidence of this problem, and fluoxetine the lowest incidence. The strength of evidence is moderate.

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

Severe adverse events.

Sexual dysfunction. Bupropion is associated with a lower incidence of sexual dysfunction than fluoxetine, paroxetine, and sertaline. The NNT to gain one additional person with high overall satisfaction of sexual functioning is 6 (95-percent CI, 4-9). In head-to-head trials, paroxetine consistently had higher rates of sexual dysfunction than comparators (fluoxetine, fluvoxamine, nefazodone, and sertraline; 16 percent vs. 6 percent). Underreporting of absolute rates of sexual dysfunction, however, is likely in these studies. Whether these findings can be extrapolated to comparisons of bupropion and paroxetine with other second-generation antidepressants is unclear. The strength of evidence is moderate.

Other severe adverse events. The existing evidence on the comparative risk for rare but severe adverse events, such as suicidality, seizures, cardiovascular events (i.e., elevated systolic and diastolic blood pressure and elevated pulse/heart rate), hyponatremia, hepatotoxicity, and serotonin syndrome, is insufficient to draw firm conclusions. The strength of evidence is low. Clinicians should keep in mind the risk of such harms during any course of treatment with a second-generation antidepressant.

Adherence. Efficacy studies do not indicate any substantial differences in adherence among second-generation antidepressants. The strength of evidence is moderate. One observational study indicated that extended-release formulations might have a better adherence rate than immediate-release medications. This finding, however, is likely attributable more to differences in dosing regimens than to differences in efficacy and harms. To what extent findings from highly controlled efficacy trials can be extrapolated to "real-world" settings remains uncertain. The evidence is insufficient to reach any conclusions about differences in adherence in effectiveness studies. The strength of evidence is low.

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

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

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

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

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

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

Remaining Issues

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

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

Major Depressive Disorder

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

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

In addition, more evidence is needed regarding the most appropriate duration of antidepressant treatment for maintaining response and remission. Such studies should also evaluate whether different formulations (i.e., controlled release vs. immediate release) lead to differences in adherence and subsequently to differences in relapse or recurrence. Additionally, although most trials maintained the dose used in acute-phase treatment throughout continuation and maintenance treatment, little is known about the effect of drug dose on the risk of relapse or recurrence. The role of other depression treatments, such as psychotherapy, vagal nerve stimulations, light therapy, and complementary medicines, as substitutes or complements to pharmaceutical management also needs to be better understood.

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

Dysthymia and Subsyndromal Depression

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

Addendum

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

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

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

Overall, 67 percent of patients achieved remission. Remission rates were 36.8 percent for Step 1, 30.6 percent for Step 2, 13.7 percent for Step 3, and 13.0 percent for Step 4. For patients achieving acceptable benefits who continued on in the 12-month followup study, relapse rates were 40.1 percent, 55.3 percent, 64.6 percent, and 71.1 percent for those achieving benefit in

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

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

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 1a. Comparative efficacy and effectiveness of second-generation antidepressants		
Major depressive disorders		
Comparative efficacy	Moderate	Results from direct and indirect comparisons indicate that no substantial differences in efficacy exist among second-generation antidepressants.
Comparative effectiveness	Moderate	Direct evidence from 1 good and 2 fair effectiveness studies and indirect evidence from efficacy trials indicate that no substantial differences in effectiveness exist among second-generation antidepressants.
Quality of life	Moderate	Consistent results from 18 studies, most of fair quality, indicate that the efficacy of second-generation antidepressants with respect to quality of life does not differ among drugs.
Onset of action	Moderate	Consistent results from 7 fair trials suggest that mirtazapine has a significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline. Whether this difference can be extrapolated to other second-generation antidepressants is unclear. Most other trials do not indicate a faster onset of action of 1 second-generation antidepressant compared with another.
Dysthymia		
Comparative efficacy	Low	No head-to head evidence exists. Findings from 5 placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	Low	One fair effectiveness study provides mixed evidence about paroxetine vs. placebo; patients older than 60 showed greater improvement on paroxetine; those younger than 50 did not show any difference.
Quality of life	No evidence	
Onset of action	No evidence	
Subsyndromal depression		
Comparative efficacy	Low	One nonrandomized, open-label trial did not detect any difference between citalopram and sertraline. Findings from 2 placebo-controlled trials were insufficient to draw conclusions.
Comparative effectiveness	No evidence	
Quality of life	No evidence	
Onset of action	No evidence	
Key Question 1b: Greater efficacy and effectiveness with previously effective medications		
Major depressive disorder	No evidence	
Dysthymia	No evidence	
Subsyndromal depression	No evidence	

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

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

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

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Melancholia		
Comparative efficacy	Low	Evidence from 2 fair head-to-head studies, 1 poor head-to-head study, and 1 fair placebo-controlled trial is insufficient to draw conclusions about treating depression in patients with coexisting insomnia. Results are inconsistent across studies.
Comparative effectiveness	No evidence	
Pain		
Comparative efficacy	Low	Evidence from 1 fair placebo-controlled study is insufficient to draw conclusions about treating depression in patients with coexisting pain. Results from head-to-head trials are not available.
Comparative effectiveness	No evidence	
Psychomotor change		
Comparative efficacy	Low	Evidence from 1 fair head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting psychomotor change. Results indicate that comparative outcomes for psychomotor retardation and psychomotor change may be different.
Comparative effectiveness	No evidence	
Somatization		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Key Question 3b: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of symptom clusters in patients with depression		
Anxiety		
Comparative efficacy	Moderate	Results from 10 fair head-to-head trials and 2 fair placebo-controlled trials suggest that no substantial differences in efficacy exist among second-generation antidepressants for treatment of accompanying anxiety symptoms.
Comparative effectiveness	No evidence	
Insomnia		
Comparative efficacy	Low	Evidence from 6 fair head-to-head trials is insufficient to draw conclusions about treating insomnia in depressed patients. Results are limited by study design; differences in outcomes are of unknown clinical significance.
Comparative effectiveness	No evidence	
Melancholia		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Pain		
Comparative efficacy	Low	Evidence from 2 head-to-head trials (1 fair, 1 poor) and 3 placebo-controlled trials is insufficient to draw conclusions about treating coexisting pain in depressed patients. Results indicate no difference in efficacy but are limited by study design.
Comparative effectiveness	No evidence	

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

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Psychomotor change		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Somatization		
Comparative efficacy	No evidence	
Comparative effectiveness	Low	Evidence from 1 open-label head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating coexisting somatization in depressed patients. Results indicate no difference in effectiveness.
Key Question 4: Comparative risk of harms		
General tolerability		
Adverse events profiles	High	Adverse events profiles are similar among second-generation antidepressants. Incidence rates of specific adverse events differ.
Nausea and vomiting	High	Meta-analysis of 15 fair studies indicates that venlafaxine has a higher rate of nausea and vomiting than SSRIs as a class.
Diarrhea	Moderate	Evidence from 15 fair studies indicates that sertraline has a higher incidence of diarrhea than bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, and venlafaxine.
Weight change	Moderate	Seven fair trials indicate that mirtazapine leads to higher weight gains than citalopram, fluoxetine, paroxetine, and sertraline.
Somnolence	Moderate	Six fair studies provide evidence that trazodone has a higher rate of somnolence than bupropion, fluoxetine, mirtazapine, paroxetine, and venlafaxine.
Discontinuation syndrome	Moderate	A good systematic review provides evidence that paroxetine and venlafaxine have the highest rates of discontinuation syndrome; fluoxetine has the lowest.
Discontinuation rates	High	Meta-analyses of efficacy trials indicate that overall discontinuation rates are similar. Venlafaxine has a higher rate of discontinuations from adverse events and a lower rate of discontinuations from lack of efficacy than SSRIs as a class.
Severe adverse events		
Sexual dysfunction	Moderate	Evidence from 5 fair trials provides evidence that bupropion causes significantly less sexual dysfunction than fluoxetine, paroxetine, and sertraline. Among SSRIs, paroxetine has the highest rates of sexual dysfunction.
Suicidality	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of suicidality.
Seizures	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of seizures. Weak evidence indicates that bupropion may have an increased risk of seizures.
Cardiovascular events	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of cardiovascular adverse events. Weak evidence indicates that venlafaxine might have an increased risk of cardiovascular adverse events.

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

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

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

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Conclusion
Race or ethnicity		
Comparative efficacy	Low	Results from 1 poor RCT indicate that efficacy does not differ substantially among second-generation antidepressants in different racial subgroups.
Comparative effectiveness	No evidence	
Comparative harms	No evidence	
Comorbidities		
Comparative efficacy	Low	One poor head-to-head trial included patients with depression and HIV/AIDS; this study indicated that efficacy does not differ substantially among second-generation antidepressants. Findings from placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	No evidence	
Comparative harms	No evidence	

[1]Strength of evidence is based on a modified version of the GRADE system; see text above.
[2]Good, fair, or poor designations relate to quality grades given to each study.
Abbreviations: RCT = randomized controlled trial; SR = sustained release; SSRI = selective serotonin reuptake inhibitor; XR = extended release.

Introduction

Background

Axis I psychiatric disorders such as depressive disorder, anxiety disorder, adjustment disorder, and premenstrual dysphoric disorders can be serious disabling illnesses. Combined, they affect approximately one in five Americans.[1] Major depressive disorder (MDD) is the most prevalent, affecting more than 16 percent (lifetime) of US adults.[2] In 2000, the US economic burden of depressive disorders was estimated to be $83.1 billion.[3] More than 30 percent of these costs were attributable to direct medical expenses.

Pharmacotherapy dominates the medical management of Axis I mood and anxiety disorders. Before the late 1980s, pharmacologic treatment was limited to tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) (with the exception of premenstrual dysphoric disorder, which historically was untreated). TCAs and MAOIs sometimes are referred to as traditional or first-generation antidepressants. They are often accompanied by multiple side effects that many patients find intolerable. For example, TCAs tend to cause anticholinergic effects including dry mouth and eyes, urinary hesitancy, and sometimes retention and constipation. In addition, TCAs have a high rate of lethality when overdose occurs; MAOIs can produce hypertensive crisis if taken along with certain foods or dietary supplements containing excessive amounts of tyramine. Thus, first-generation antidepressants are no longer agents of choice in many circumstances.

Newer treatments include selective serotonin reuptake inhibitors (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs), and other second-generation drugs. The first of the second-generation drugs was introduced to the US market in 1982, when trazodone was approved for treatment of patients with MDD. In 1987, the US Food and Drug Administration (FDA) approved the first SSRI, fluoxetine. Since then, five other SSRIs have been introduced: sertraline (1991), paroxetine (1992), citalopram (1999), fluvoxamine (2000), and escitalopram (2002).

Two other second-generation antidepressant drugs, trazodone and nefazodone, also function as serotonin reuptake inhibitors, but they possess additional serotonin antagonist properties. Trazodone, which was first synthesized in 1966, appears to produce its primary effect by selectively inhibiting serotonin reuptake, but it also causes adrenoreceptor subsensitivity and induces significant changes in 5-hydroxytryptamine (5-HT) presynaptic receptor adrenoreceptors. Although approved for MDD, trazodone commonly is used as a sedative to complement newer stimulating antidepressants. In 1994, the FDA approved nefazodone, which is essentially an SSRI with additional 5-HT2 and 5-HT3 antagonist properties.

The SNRIs were first introduced to the market in 1993 with the approval of venlafaxine. Mirtazapine, a drug that acts centrally on adrenergic autoreceptors, was added to the therapeutic arsenal in 1996.[4] Duloxetine, a selective serotonin and norepinephrine reuptake inhibitor (SSNRI), was approved for the treatment of MDD and diabetic peripheral neuropathic pain in 2004.

The mechanism of action of most second-generation antidepressants is poorly understood. In general, these drugs work through their effect on prominent neurotransmitters in the central

nervous system. Bupropion is a relatively weak inhibitor of the neuronal uptake of norepinephrine, serotonin, and dopamine; its primary mechanism of action is believed to be dopaminergic and noradrenergic. The SSRIs (citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline) act by selectively inhibiting the reuptake of serotonin 5-HT at the presynaptic neuronal membrane. Reuptake inhibition has the effect of increasing the levels of serotonin made available to improve the transmission of neural signals at the synapse. The SNRIs (venlafaxine) are potent inhibitors of serotonin and norepinephrine reuptake and weak inhibitors of dopamine reuptake. Mirtazapine, sometimes characterized as an SNRI, is believed to enhance central noradrenergic and serotonergic activity as a 5-HT2 and 5-HT3 receptor antagonist. Trazodone inhibits neuronal uptake of serotonin. At low doses, it appears to act as a serotonin antagonist and at higher doses as an agonist.[5,6] Nefazodone is believed to inhibit neuronal uptake of serotonin and norepinephrine. Preclinical studies of duloxetine suggest that it is a potent inhibitor of neuronal serotonin and norepinephrine reuptake and a less potent inhibitor of dopamine reuptake.

Except for fluvoxamine, which is approved only for the treatment of obsessive-compulsive disorder (OCD), all second-generation antidepressants are approved for the treatment of MDD. Table 1 summarizes the newer products that are available in the United States by mechanism of action; it shows names, all dosage forms, therapeutic class, and FDA-approved (labeled) uses. The second-generation antidepressants have established a prominent role in the US pharmaceutical market. In 2003, the antidepressant class, including SSRIs and SNRIs, ranked third in US prescription sales among all drug therapy classes.[7] The serotonergic class dominates this market, accounting for 57.6 percent of market share in 2002.[7] Prescription drug spending for these products is not anticipated to decline until 2009, when the patents for leading brands will expire.

Compared with the first-generation antidepressants, the SSRIs and other second-generation antidepressant have comparable efficacy and comparable or better side effect profiles.[8,9] However, comparative differences in efficacy, tolerability, and safety are not well defined among the second-generation drugs. The tremendous volume and large variability in the quality of evidence to support use of these products makes it difficult for clinicians, patients, and others to make evidence-based decisions.

Purpose of this Report

The purpose of this review is to help policymakers and clinicians make informed choices about the use of SSRIs and newer antidepressants. Given the prominent role of drug therapy in psychiatric disease and the prevalent use of these drugs, our goal is to summarize comparative data on the efficacy, effectiveness, and harms of newer antidepressants: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, paroxetine, sertraline, trazodone, nefazodone, and venlafaxine. We examine the role of these agents in treating patients with depressive syndrome, including MDD, dysthymic disorder, and subsyndromal depressive disorders. We focus this review on these disorders in adult populations.

This report extends prior analyses by addressing two areas that are relevant for clinicians and policymakers but that previous reports have not covered. First, we consider treatment in the continuation and maintenance phases of depression, not simply the acute phase of treatment (see

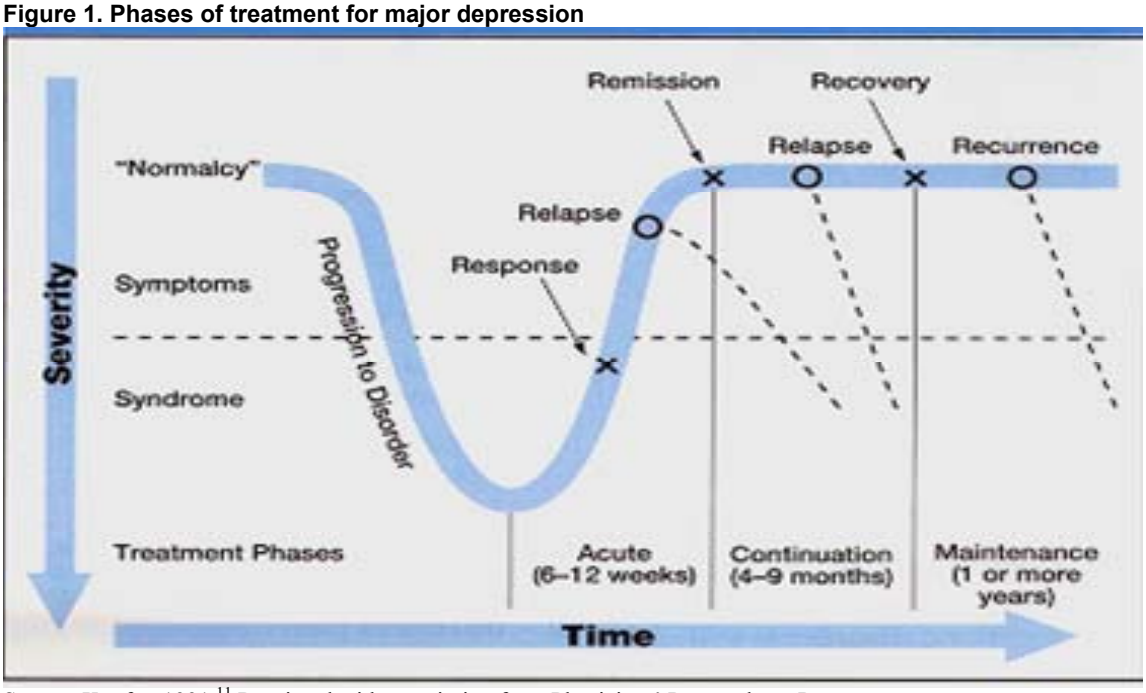
Figure 1). Previous estimates suggest that continuing treatment beyond the acute phase can reduce the odds of relapse by 70 percent.[10] However, most reports have been limited to outcomes in the acute phase of management, i.e., the initial part of treatment during which the treatment goal is eradication of the depressive symptoms to achieve remission.

Table 1. Second-generation antidepressants approved for use in the United States

Generic Name	US Trade Name*	Dosage Forms	Therapeutic Classification	Labeled Uses†
Bupropion‡	Wellbutrin®; Wellbutrin SR®; Wellbutrin XL®; Zyban®	75, 100 mg tabs; 50, 100, 150, 200 mg SR tabs 150, 300 mg XL tabs	Other	MDD
Citalopram‡	Celexa®	10, 20, 40 mg tabs; 1, 2 mg/ml solution	SSRI	MDD
Duloxetine	Cymbalta®	20, 30, 60 mg caps	Other	MDD; Neuropathic pain
Escitalopram	Lexapro®§	10, 20 mg tabs 1 mg/ml solution	SSRI	MDD; GAD
Fluoxetine‡	Prozac®; Prozac Weekly®; Sarafem®	10, 20, 40 mg caps; 10 mg tabs; 4 mg/ml solution; 90 mg pellets (weekly)	SSRI	MDD; OCD; PMDD; Panic disorder
Fluvoxamine‡	Luvox®‖	25, 50, 100 mg tabs	SSRI	OCD
Mirtazapine‡	Remeron®	15, 30, 45 mg tabs; 15, 30, 45 mg orally disintegrating tabs	Other	MDD
Nefazodone‡	Serzone®‖	50, 100, 150, 200, 250 mg tabs	Other	MDD
Paroxetine‡	Paxil®; Paxil CR®	10, 20, 30, 40 mg tabs; 2 mg/ml solution; 12.5, 25, 37.5 mg CR tabs	SSRI	MDD; OCD; Panic disorder; Social anxiety disorder; GAD; PTSD; PMDD¶
Sertraline‡	Zoloft®	25, 50, 100 mg tabs; 20 mg/ml solution	SSRI	MDD; OCD; Panic disorder; PTSD; PMDD; Social anxiety disorder
Trazodone‡	Desyrel®	50, 100, 150, 300 mg tabs	Other	MDD
Venlafaxine	Effexor®; Effexor XR®	25, 37.5, 50, 75, 100 mg tabs; 37.5, 75, 150 mg XR caps	SNRI	MDD; GAD;** Social anxiety disorder**

* CR, SR, XL, and XR are registered trademarks referring to controlled, sustained, or extended-release dosage forms.
† GAD, generalized anxiety disorder; MDD, major depressive disorder; OCD, obsessive-compulsive disorder; PMDD, premenstrual dysphoric disorder; PTSD, post-traumatic stress disorder.
‡ Generic available for some dosage forms.
§ Lexapro was denied approval for social anxiety disorder March 30, 2005.
‖ Brand-name product no longer available in the US.
¶ Only Paxil CR® (not Paxil®) is approved for the treatment of PMDD.
** Only Effexor XR® (not Effexor®) is approved for the treatment of GAD and social anxiety disorder.

Figure 1. Phases of treatment for major depression



Source: Kupfer, 1991.[11] Reprinted with permission from Physicians' Postgraduate Press.

We consider all three phases of depression management:

- acute phase, usually 6 to 12 weeks in length;
- continuation phase, during which the treatment goal is continued absence of depressive symptoms for an additional 4 to 9 months such that the patient's episode can be considered completely resolved; and
- maintenance phase, the frequently multi-year period during which the treatment goal is preventing the recurrence of a new, distinct episode.

This categorization allows us to make the clinically relevant distinction between relapse and recurrence. We define relapse as the return of depressive symptoms during the acute or continuation phases, so it is considered part of the same depressive episode. We define recurrence as the return of depressive symptoms during the maintenance phase, so it is considered a new, distinct episode.

This distinction is critical to determining the long-term treatment plan. If an individual has a single episode of MDD that has resolved, treatment recommendations may or may not include continued medication treatment. If, however, an individual has a diagnosis of recurrent MDD, the recommendation for continued treatment may be years.[12,13] In addition, this categorization can frame decisions about depression management into best treatments for immediate resolution of depressive symptoms (acute phase) and those best for ongoing management once symptoms have resolved (continuation and maintenance phases). Of note, the latter two phases involve a time period of much greater duration than the first one.

Second, we review the data addressing whether the presence of accompanying symptoms, such as anxiety and insomnia, might affect outcomes. For example, MDD is frequently associated with concurrent anxiety. If certain antidepressants can more successfully treat such a depression than other agents, or if they can mitigate the specific concurrent anxiety symptoms, these agents might be preferred choices. Such data could guide clinicians on how better to target antidepressant selection and steer policymakers toward the best available agents.

Table 1 (above) and Table 2 provide detailed information on second-generation agents approved for use in the United States. Table 2 shows trade names, usual (recommended) daily doses, and frequency.

Table 2. Usual dosing range and frequency of administration for adults

Generic Name	US Trade Name*	Usual Daily Dosing Range	Frequency
Bupropion	Wellbutrin®	100-450 mg	Three times daily
	Wellbutrin SR®	150-400 mg	Twice daily
	Wellbutrin XL®	150-450 mg	Once daily
Citalopram	Celexa®	20-60 mg	Once daily
Duloxetine	Cymbalta®	40-60 mg	Once or twice daily
Escitalopram	Lexapro®	10-20 mg	Once daily
Fluoxetine	Prozac®	10-80 mg	Once or twice daily
	Prozac Weekly®	90 mg (weekly)	Once weekly
	Sarafem®	20 mg	Once daily†
Fluvoxamine	Luvox®§	50-300 mg	Once or twice daily
Mirtazapine	Remeron®	15-45 mg	Once daily
Nefazodone‡	Serzone®§	200-600 mg	Twice daily
Paroxetine	Paxil®	10-60 mg	Once daily
	Paxil CR®	12.5-75 mg	Once daily
Sertraline	Zoloft®	25-200 mg	Once daily
Trazodone	Desyrel®	150-400 mg	Three times daily
Venlafaxine	Effexor®	75-375 mg	Two to three times daily
	Effexor XR®	75-225 mg	Once daily

* CR, SR, XL, and XR are registered trademarks referring to controlled-, sustained-, or extended-release dosage forms.
† Sarafem is marketed for the treatment of premenstrual dysphoric disorder (PMDD); dosing may be continuous or intermittent.
‡Branded product withdrawn from the US market effective June 14, 2004.
§ Brand-name product no longer available in the US.

Scope and Key Questions

This review compares the efficacy, effectiveness, and harms of second-generation antidepressant medications. To that end, we address the following key questions:

1a. For adults with major depressive disorder (MDD), dysthymia, or subsyndromal depressive disorders, do commonly used medications for depression differ in efficacy or effectiveness in treating depressive symptoms?

1b. If a patient has responded to one agent in the past, is that agent better than current alternatives at treating depressive symptoms?

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

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

3. Do medications or combinations of medications (including tricyclics in combination) used to treat depression differ in their efficacy or effectiveness for treating accompanying symptoms, such as anxiety, insomnia, and neurovegetative symptoms? This question focuses on accompanying neurovegetative (physical) symptoms of depression (such as disturbances in sleep, appetite, or motor activity; or symptoms of fatigue or pain). These symptom clusters are in contrast to psychological or cognitive symptoms, such as worthlessness, hopelessness, excessive guilt, and suicidal ideation. For patients presenting with these accompanying symptom clusters, two treatment outcomes are relevant: the effect on the depressive disorder overall, and the effect on the particular accompanying symptoms. Consequently, we further divide this question into two subquestions:

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

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

4. For adults with a depressive syndrome, do commonly used antidepressants differ in safety, adverse events, or adherence? Adverse effects of interest include but are not limited to nausea, diarrhea, headache, tremor, daytime sedation, decreased libido, failure to achieve orgasm, nervousness, insomnia, and more serious events including suicide.

5. How do the efficacy, effectiveness, or harms of treatment with antidepressants for a depressive syndrome differ for the following subpopulations?

    - elderly or very elderly patients;
    - other demographic groups (defined by age, ethnic or racial groups, and sex);
    - patients with medical comorbidities (e.g., ischemic heart disease, cancer);
    - patients with psychiatric and behavioral comorbidities (e.g., substance abuse disorders); and
    - patients taking other medications.

Throughout this report, we highlight *effectiveness* studies conducted in primary care or office-based settings that use less stringent eligibility criteria, assess health outcomes, and have longer followup periods than most *efficacy* studies.[14] We deemed studies that met at least six of seven predefined criteria as effectiveness studies (Table 3). Their results are more applicable to the average patient than results from highly selected populations in efficacy studies.

Table 3. Criteria for effectiveness studies

Criteria	Relevance to Treatment of Depressive Disorders
Study population	Primary care population
Less stringent eligibility criteria	Determine case by case
Health outcomes	Response, remission, quality of life, functional capacity, hospitalization
Clinically relevant treatment modalities	$\geq$ 8 weeks study duration; flexible dose design; physician-based diagnosis
Assessment of adverse events	Always
Adequate sample size to assess a minimally important difference from a patient perspective	N $\geq$ 150
Intention-to-treat analysis	Always

For each key question, we evaluated specific outcome measures (where appropriate), as reported in Table 4. For efficacy and effectiveness, we focused on head-to-head trials comparing one second-generation antidepressant with another. When sufficient head-to-head evidence was unavailable, we evaluated placebo-controlled evidence. Finally, we included observational studies to assess relapse or recurrence prevention, second-line treatment, and safety and tolerability.

Table 4. Outcome measures and study eligibility criteria

Key Question Outcomes of Interest and Specific Measures	Study Eligibility Criteria
Key Questions 1, 3, and 5: Efficacy and effectiveness Response Remission Speed of response/remission Relapse Quality of life Functional capacity Hospitalization	**Study design** • Head-to-head, double-blind, RCTs • High-quality meta-analyses • When sufficient evidence is not available for direct head-to-head comparisons: double-blind, placebo-controlled RCTs **Minimum study duration** • 6 weeks **Study population** • Adult inpatients and outpatients with major depressive disorder, dysthymia, or subsyndromal depression **Sample size** • For quantitative analysis: no limit • For qualitative analysis: n ≥ 40
Key Question 2a: Maintenance of remission	**Study design** • Head-to-head, double-blind, RCTs • High-quality meta-analyses • High-quality, controlled observational studies **Minimum study duration** • For all studies: 3 months **Study population** • Adult inpatients and outpatients with a history of depressive illnesses currently in remission **Sample size** • For RCTs: no limit • For observational studies: n ≥ 100

RCT, randomized controlled trial.

Table 4. Outcome measures and study eligibility criteria (continued)

Key Question Outcomes of Interest and Specific Measures	Study Eligibility Criteria
Key Question 2b: **Response and remission for recurrent depression**	**Study design** • Head-to-head, double-blind, RCTs • High-quality meta-analyses • High-quality, controlled observational studies **Minimum study duration** • For RCTs: 6 weeks • For observational studies: 3 months **Study population** • Adult inpatients and outpatients with recurrent depression **Sample size** • For RCTs: 1. For quantitative analysis: no limit 2. For qualitative analysis: n ≥ 40 • For observational studies: n ≥ 100
Key Question 4: **Safety and tolerability:** • Overall adverse events • Withdrawals because of adverse events • Serious adverse events • Specific adverse events or withdrawals because of specific adverse events, including: 3. hyponatremia 4. seizures 5. suicide 6. hepatoxicity 7. weight gain 8. gastrointestinal symptoms 9. sexual side effects **10.** others	**Study design** • Head-to-head, double-blind, RCTs • High-quality meta-analyses • Observational studies (cohort studies, case-control studies, large database reviews) **Minimum study duration** • For RCTs: 6 weeks For observational studies: 3 months **Study population** • Adult inpatients and outpatients with major depressive disorder, dysthymia, or subsyndromal depression **Sample size** • For RCTs: 11. For quantitative analysis: no limit 12. For qualitative analysis: n ≥ 40 • For observational studies: n ≥ 100

Appendix A lists our peer reviewers. Appendices B-I pertain to aspects of our results.

Methods

Topic Development

The topic of this report and preliminary key questions arose through a public process involving the public, the Scientific Resource Center (www.effectivehealthcare.ahrq.gov/aboutUs/contact.cfm) for the Effective Health Care program of the Agency for Healthcare Research and Quality (AHRQ) (www.effectivehealthcare.ahrq.gov), and various stakeholder groups (www.effectivehealthcare.ahrq.gov/aboutUs/stakeholder.cfm). Investigators from the RTI International-University of North Carolina Evidence-based Practice Center (RTI-UNC EPC) then refined the questions in consultation with AHRQ, the Scientific Resource Center, and a Technical Expert Panel (TEP).

Literature Search

To identify articles relevant to each key question (Appendix B), we searched MEDLINE®, Embase, The Cochrane Library, PsychLit, and the International Pharmaceutical Abstracts. We used either Medical Subject Headings (MeSH or MH) as search terms when available or key words when appropriate. We combined terms for selected indications (major depressive disorder [MDD], dysthymia, minor depression, subsyndromal depressive disorder), drug interactions, and adverse events with a list of 12 specific second-generation antidepressants (bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine). We limited the electronic searches to "human" and "English language." Sources were searched from 1980 to 2006 (February) to capture literature relevant to the scope of our topic.

We used the National Library of Medicine (NLM) publication type tags to identify reviews, randomized controlled trials (RCTs), and meta-analyses. We also manually searched reference lists of pertinent review articles and letters to the editor. We imported all citations into an electronic database (EndNote 9.0). Additionally, we handsearched the Center for Drug Evaluation and Research (CDER) database to identify unpublished research submitted to the US Food and Drug Administration (FDA).

The Scientific Resource Center contacted pharmaceutical manufacturers and invited them to submit dossiers, including citations. We received dossiers from three pharmaceutical companies (Eli Lilly, GlaxoSmithKline, and Wyeth).

Our searches found 1,967 citations, unduplicated across databases. Additionally, we detected 129 articles from manually reviewing the reference lists of pertinent review articles. Three other studies came from pharmaceutical dossiers. The total number of citations in our database was 2,099.

Study Selection

We developed eligibility criteria with respect to study design or duration, patient population, interventions, outcomes, and comparisons to antidepressant medications outside our scope of interest, as described in Table 4 (in the introduction). Two persons independently reviewed abstracts. If both reviewers agreed that the trial did not meet eligibility criteria, we excluded it. We obtained the full text of all remaining articles and used the same eligibility criteria to determine which, if any, to exclude at this stage.

For this review, results from well-conducted, valid head-to-head trials provide the strongest evidence to compare drugs with respect to efficacy, effectiveness, and harms. We defined head-to-head trials as those comparing one second-generation antidepressant with another. RCTs of at least 6 weeks' duration and an adult study population with a sample size of at least 40 participants were eligible for inclusion.

We did not examine placebo-controlled trials in detail if head-to-head trials were available. If no head-to-head evidence was published, we reviewed placebo-controlled trials to provide an overview of efficacy. For harms (i.e., evidence pertaining to safety, tolerability, and adverse events), we examined data from both experimental and observational studies. We included observational studies with large sample sizes ($\geq$100 patients), lasting at least 3 months, that reported an outcome of interest.

Initially, we reviewed studies with health outcomes as primary outcome measures. Outcomes for efficacy or effectiveness, for example, were quality of life, relapse, functional capacity, and hospitalization. If no study measuring health outcomes was available for a particular indication or population subgroup, we included intermediate outcomes (e.g., changes in depression scores). We reviewed response and remission when based on changes in depression scores as proxies for health outcomes (e.g., 50 percent improvement of depression scores for response). For harms (throughout this report we use "harms" as a summary term for adverse events and unwanted effects, as suggested by the CONSORT [Consolidated Standards of Reporting Trials] statement), we looked for both overall and specific outcomes ranging in severity (e.g., suicide, sexual side effects, hyponatremia, weight change, seizures, gastrointestinal symptoms, discontinuation syndrome), withdrawals attributable to adverse events, and drug interactions.

We included meta-analyses in our evidence report if we found them to be relevant for a key question and of good or fair methodological quality (based on the QUORUM statement[15]). We did not review individual studies if they were included in a high-quality meta-analysis. We excluded meta-analyses that were not based on a comprehensive systematic literature search or did not maintain the units of the studies in their statistical analyses. We checked our database to guarantee that our literature search had detected trials included in any meta-analyses that we discarded, and we then obtained any missing articles.

Data Extraction

We designed and used a structured data abstraction form to ensure consistency of appraisal for each study. Trained reviewers abstracted data from each study and assigned an initial quality

rating. A senior reviewer read each abstracted article, evaluated the completeness of the data abstraction, and confirmed the quality rating.

We abstracted the following data from included trials: study design, eligibility criteria, intervention (drugs, dose, duration), additional medications allowed, methods of outcome assessment, population characteristics (such as age, sex, race or ethnicity, or comorbid anxiety), sample size, loss to followup, withdrawals because of adverse events, results, and adverse events reported. We recorded intention-to-treat results (ITT) if available. All data abstraction employed SRS 3.0, TrialStat™ Corporation.

Quality Assessment

To assess the quality (internal validity) of trials, we used predefined criteria based on those developed by the US Preventive Services Task Force (ratings: good, fair, poor)[16] and the National Health Service Centre for Reviews and Dissemination.[17] Elements of quality assessment included, among others, randomization and allocation concealment, similarity of compared groups at baseline, use of ITT analysis (i.e., all patients are analyzed as randomized with missing values imputed), and overall and differential loss to followup.

In general terms, a "good" study has the least bias and results are considered to be valid. A "fair" study is susceptible to some bias, but probably not sufficient to invalidate its results. The fair quality category is likely to be broad, so studies with this rating will vary in their strengths and weaknesses. A "poor" rating indicates significant bias (stemming from, e.g., serious errors in design, analysis reporting large amounts of missing information, or discrepancies in reporting) that may invalidate the study's results.

To assess the quality of observational studies, we used criteria outlined by Deeks et al.[18] Items assessed included selection of cases or cohorts and controls, adjustment for confounders, methods of outcomes assessment, length of followup, and statistical analysis.

Two independent reviewers assigned quality ratings. They resolved any disagreements by discussion and consensus or by consulting a third, independent party.

Studies that met all criteria were rated good quality. The majority of studies received a quality rating of fair. This category includes studies that presumably fulfilled all quality criteria but did not report their methods to an extent that answered all our questions. Time constraints precluded our contacting study authors for clarification of methodological questions. Thus, the fair quality category includes studies with quite different strengths and weaknesses. Studies that had a fatal flaw in one or more categories were rated poor quality and, generally, excluded from our analyses. If no other evidence on an outcome of interest was available, we comment on findings from poor studies.

In addition to internal and external validity, we assessed the comparability of dosages. Because we could not find any clear definitions about equivalence of dosages among second-generation antidepressants in the published literature, we developed a roster of low, medium, and high dosages for each drug, which is outlined in Table 5. This classification, based on the interquartile dosing range, does not indicate dosing equivalence. We used this roster to detect gross inequalities in dosing that could affect comparative efficacy and effectiveness.

Applicability Assessment

Throughout this report, we highlight *effectiveness* studies conducted in primary care or office-based settings that use less stringent eligibility criteria, assess health outcomes, and have longer followup periods than most *efficacy* studies. The results of effectiveness studies are more applicable to the spectrum of patients that will use a drug, have a test, or undergo a procedure than results from highly selected populations in efficacy studies. We used criteria proposed by Gartlehner et al. to distinguish effectiveness from efficacy trials.[19] These criteria assess seven categories: primary care population, eligibility criteria, outcome measures, study duration and intervention modalities, adverse events assessment, sample size, and ITT analysis.

Table 5. Dosing classification based on lower and upper dosing range quartiles

Drug	Range	Low	Medium	High
Bupropion	300-450 mg/d	< 337.5	337.5-412.5	> 412.5
Citalopram	20-60 mg/d	< 30	30-50	> 50
Duloxetine	60-100 mg/d	< 70	70-90	> 90
Escitalopram	10-30 mg/d	< 15	15-25	> 25
Fluoxetine	20-60 mg/d	< 30	30-50	> 50
Fluvoxamine	50-150 mg/d	< 75	75-125	> 125
Mirtazapine	15-45 mg/d	< 22.5	22.5-37.5	> 37.5
Nefazodone	300-600 mg/d	< 375	375-525	> 525
Paroxetine	20-60 mg/d	< 30	30-50	> 50
Sertraline	50-150 mg/d	< 75	75-125	> 125
Trazodone	300-600 mg/d	< 375	375-525	> 525
Venlafaxine	125-250 mg/d	< 156.25	156.25-218.75	> 218.75

Rating Strength of a Body of Evidence

We rated the strength of the available evidence in a three-part hierarchy based on an approach devised by the GRADE working group.[20] Developed to grade the quality of evidence and the strength of recommendations, this approach incorporates four key elements: study design, study quality, consistency, and directness. As shown in Table 6, we used three grades: high, moderate, and low (combining the GRADE category of very low with low).[21] Gradings reflect the strength of the body of evidence to answer key questions on the comparative efficacy, effectiveness, and harms of second-generation antidepressants. Gradings do not refer to the general efficacy or effectiveness.

Table 6. Definitions of the grades of the overall quality of evidence

High	Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Further research is likely to have an important impact on our confidence in the estimate of the effect and may change the estimate.
Low	Further research is very likely to have an important impact on our confidence in the estimate of the effect and is likely to change the estimate.

Source: Adapted from the GRADE working group.[20]

This approach does not incorporate other factors that might be relevant to assess reliably the comparative efficacy, effectiveness, and harms, such as funding sources and comparable dosing. We have assessed these additional factors and highlighted inequalities that could potentially bias our assessments (e.g., all studies funded by the same manufacturer).

Data Synthesis

Throughout this report we synthesized the literature qualitatively. If data were sufficient, we augmented findings with quantitative analyses. We conducted meta-analyses of data for head-to-head comparisons for trials that were fairly homogenous in study populations and outcome assessments. For efficacy, we used two outcome measures:

1. The relative risk (RR) of being a responder (more than 50 percent improvement from baseline) on the Hamilton Rating Scale for Depression (HAM-D) or the Montgomery-Asberg Depression Rating Scale (MADRS) at study endpoint.

2. The weighted mean difference of changes on a specific depression rating scale (HAM-D or MADRS). We chose this outcome measure to have an estimate of the actual difference in effect sizes between treatments.

For each meta-analysis, we conducted a test of heterogeneity (I^2 statistic) and applied both a random and a fixed effects model. We report the results from random effects models because, in all meta-analyses, the results from random and fixed effects models were very similar. If the RR was statistically significant, we then conducted a meta-analysis of the risk differences to calculate the number needed to treat (NNT) on the pooled risk difference.

We assessed publication bias using funnel plots and Kendell's tests. However, given the small number of component studies in our meta-analyses, these tests have low sensitivity to detect publication bias.

If fewer than three head-to-head trials were available for any drug comparison, we computed indirect comparisons. Evidence suggests that indirect comparisons agree with head-to-head trials if component studies are similar and treatment effects are expected to be consistent in patients in different trials.[22]

For indirect comparisons we employed two statistical approaches:

1. We conducted meta-regressions of good or fair placebo-controlled trials using individual drugs as covariates. We also attempted to assess the influence of disease

severity, concomitant anxiety, and dosing on our findings. Data, however, were insufficient to use these factors as covariates in the meta-regression models.

2. When the number of placebo-controlled trials was insufficient to conduct meta-regressions, we used network meta-analyses.[23] Network meta-analyses allow the use of studies with multiple common comparators. Therefore, we could include both placebo-controlled and active-controlled studies, increasing the precision of results. All statistical analyses used StatsDirect Ltd. version 2.4.5, and STATA 9.1.

Results

Overview of All Key Questions

We identified 2,099 citations from searches and reviews of reference lists. Figure 2 documents the disposition of the 293 articles in this review, working from 884 articles retrieved for full review, 66 included for background, and 525 excluded at this stage (Appendix C). One study of interest could not be retrieved after multiple attempts. We included 293 published articles reporting on 187 studies of good or fair quality: 89 head-to-head randomized controlled trials (RCTs) (94 articles), 57 placebo-controlled RCTs (63 articles), 7 articles on meta-analyses or systematic reviews (7 articles), 20 observational studies (27 articles), and 14 studies of other design (16 articles). We incorporated data from 24 additional placebo-controlled studies for meta-regression only. Evidence tables for included studies, by key question, can be found in Appendix D.

Reasons for exclusion were based on eligibility criteria or methodological criteria. We excluded 62 studies that originally met eligibility criteria but were later rated as poor quality for internal validity (Appendix E). The two main reasons for rating as poor of RCTs were high loss to followup (more than 40 percent) and lack of intention-to-treat (ITT) analysis. Among meta-analyses, lack of a systematic literature search was the main reason for exclusion; this problem leads to a selected spectrum of trials and subsequently to biased results.

Most efficacy trials were not powered to establish a greater efficacy of a particular drug over another. Therefore, we report differences in effect sizes for individual studies, even if they are not statistically significant.

Of 187 included studies, 126 (67.4 percent) were financially supported by pharmaceutical companies; 17 (9.2 percent) were funded by governmental agencies or independent funds. For 44 (23.5 percent) of included studies, we could not determine the funding source.

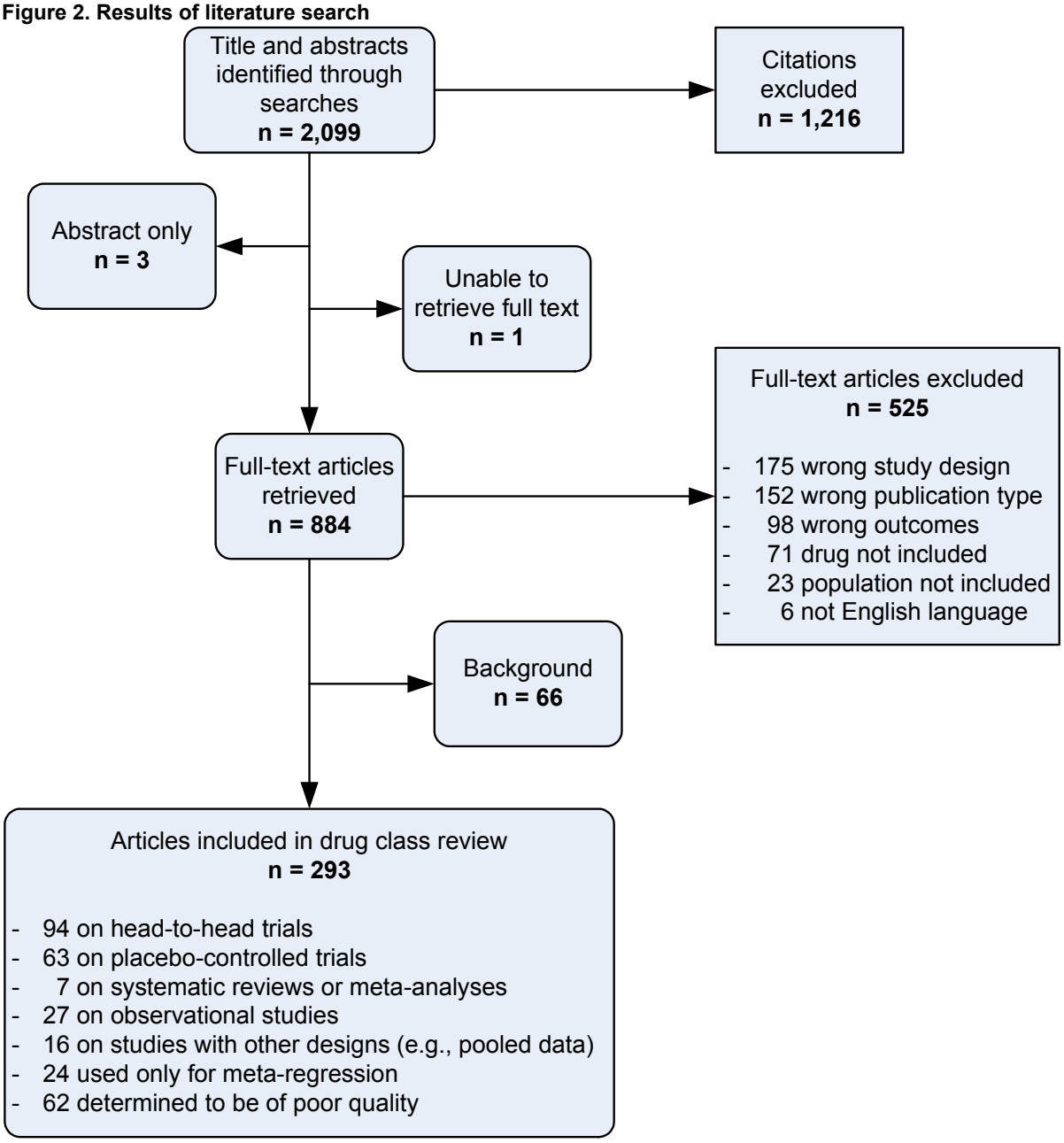
Studies reviewed for this report employed a notable array of diagnostic scales and health status or quality of life instruments. Most were pertinent to depressive and other disorders considered in this report, but some are considered more generic instruments that assess, e.g., health-related quality of life. Table 7 lists abbreviations of diagnostic scales and health status or quality of life instruments encountered in this literature.

Key Question 1: Efficacy or effectiveness in treating depressive disorders and symptoms

1a. Do commonly used medications for depression differ in efficacy or effectiveness in treating depressive symptoms in adults with major depressive disorder (MDD), dysthymia, or subsyndromal depressive disorders?

1b. If a patient has responded to one agent in the past, is that agent better than current alternatives at treating depressive symptoms?

Figure 2. Results of literature search



Major Depressive Disorder (MDD): Overview

The following second-generation antidepressants are currently approved by the US Food and Drug Administration (FDA) for the treatment of depressive disorders in adults: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, mirtazapine, nefazodone, paroxetine, sertraline,

trazodone, and venlafaxine. Fluvoxamine has not been approved for the treatment of MDD but was included on the list of medications of interest that Agency for Healthcare Research and Quality (AHRQ) provided to us for this review.

Table 7. Abbreviations and full names of diagnostic scales and other instruments

Abbreviation	Full Name of Instrument
BDI	Beck Depression Inventory
Beck's SSI	Scale for Suicide Ideation
BIMT	Blessed Information and Memory Test
BPI	Brief Pain Inventory
BQOL	Battelle Quality of Life Measure
BQOLS	Battelle Quality of Life Scale
BSI	Brief Symptom Inventory of Depression
CAS	Clinical Anxiety Scale
CES-D	Center for Epidemiological Studies-Depression Scale
CGI	Clinical Global Impressions
CGI–I	Clinical Global Impressions Improvement Scale
CGI–S	Clinical Global Impressions Severity Scale
CLAS	Clifton Assessment Schedule
DESS	Discontinuation Emergent Signs and Symptoms Checklist
FSCL	Fatigue Symptoms Checklist
FSQ	Functional Status Questionnaire
GAF-S	Global Assessment of Functioning Scale
HAD-A	Hospital Anxiety and Depression Rating Scale
HAM–A	Hamilton Rating Scale for Anxiety
HAM–D	Hamilton Rating Scale for Depression
HADRS	Hamilton Depression Rating Scale
HSCL-D20	Hopkins Symptom Checklist - Depression
IDAS	Irritability, Depression, and Anxiety Scale
IDS C	Inventory for Depressive Symptomatology - Clinician Rated
IDS SR	Inventory for Depressive Symptomatology - Self Rated
MADRS	Montgomery-Asberg Depression Rating Scale
MAF	Multidimensional Assessment of Fatigue
MMSE	Mini Mental State Examination
PGI-I	Patient Global Impression of Improvement
POMS-FI	Profile of Mood States Fatigue/Inertia Subscale
PRSexDQ	Psychotropic-Related Sexual Dysfunction Questionnaire
QLDS	Quality of Life in Depression Scale
Q-LES-Q, QLSQ	Quality of Life Enjoyment and Satisfaction Questionnaire
SCL 56	Hopkins Symptom Checklist- 56 item version
SF-36	Medical Outcomes Study Health Survey – Short Form 36
SIP	Sickness Impact Profile
SLT	Shopping List Task
VAS	Visual Analogue Scale
UKU-SES	Utvalg for Kliniske Undersogelse Side Effect Scale

In all, 72 RCTs (reported in 74 articles) compared the efficacy or effectiveness of one second-generation antidepressant with another for treating patients with MDD. Details can be found in Evidence Table 1, Appendix D.

Table 8 provides selected information on all these studies; they are grouped according to the main classes compared—selective serotonin reuptake inhibitors (SSRI) vs. SSRI, SSRI vs. selective serotonin and norepinephrine reuptake inhibitor (SSNRI) and SNRI; SSRI vs. other second-generation antidepressants—and then listed alphabetically by the specific drugs compared. Most subjects were younger than 60 years; 10 trials were conducted in populations of 60 years or older. In the text below, studies are of fair quality unless otherwise specified.

In general, studies enrolled patients according to a criteria-based diagnosis of MDD relating to the Diagnostic and Statistical Manual of Mental Disorders (DSM, either revised third edition or fourth edition [DSM-III-R, DSM-IV]) and a predefined cutoff point of a widely used depression scale (i.e., Hamilton Rating Scale for Depression [HAM-D] = 18 or Montgomery-Asberg Depression Rating Scale [MADRS] = 19). Most patients had moderate to severe depression as measured by a variety of scales. Most studies excluded patients who had additional Axis I disorders, high suicidal risk, or progressive medical diseases or who used psychotherapy, electroconvulsive therapy, or psychotropic medications.

Of 66 possible comparisons of included second-generation antidepressants (51 involving SSRIs, 15 more involving non-SSRI agents), we found direct head-to-head evidence for only 35 comparisons (30 and 5, respectively). Tables 9 and 10 depict possible comparisons and the numbers of available head-to-head trials for each comparison (shown in italics). For those with fewer than three head-to-head trials, we conducted indirect comparisons. Appendix E presents placebo-controlled studies included for indirect comparisons; Appendix F lists studies excluded from indirect comparisons because of poor internal validity.

Investigators rarely assessed quality of life and functional capacity; if they did, they typically considered these as only secondary outcomes. Most studies employed both physician-rated scales (e.g., HAM-D, MADRS, Clinical Global Impressions Scale [CGI]) and patient-rated scales (e.g., Hospital Anxiety and Depression Rating Scale [HAD-A], Battelle Quality of Life Scale [BQOLS]). All studies used physician-rated scales to assess the main outcome measures.

In the majority of studies, the primary endpoints were either changes from baseline or rates of response or remission on investigator-rated diagnostic depression scales such as the HAM-D or MADRS. Changes on such diagnostic depression scales are generally viewed as intermediate outcomes rather than health outcomes, and they are not always reliably related to changes in health outcomes. Response or remission, even when deducted from such a scale (e.g., response is defined as a 50 percent improvement of scores on HAM-D or MADRS), could be seen as proxies to health outcomes. Therefore, we focused on differences in response or remission rates rather than differences in changes of scores.

Most studies received a fair rating for internal validity. The applicability of the results was hard to determine and might often be limited. Most trials (70 percent) were of short (6 weeks to 8 weeks) or medium (9 weeks to 11 weeks) duration; 30 percent reported followup of 12 weeks or more. Short-term studies may be limited in their ability to account appropriately for response rates and long-term adverse events. In addition, reviewed studies were conducted over a time span of 2 decades. Therefore, study populations might vary with respect to cotreatment, prior exposures to other second-generation antidepressants, and other factors.

Trial reporting was often incomplete. Most articles did not report the method of randomization or allocation concealment. Last-observation-carried-forward methods (or LOCF

analysis, which means that the last observed measurement serves as the substitute for missing values because of the drop out of patients at different time points) were a frequent approach to

Table 8. Study characteristics, response and remission rates, and quality ratings of studies in adults with major depressive disorder

Study	N	Duration	Comparison and Dose (mg/d)	Response (%) and Significance Level	Remission (%) and Significance Level	Quality Rating
SSRIs vs. SSRIs						
Burke et al., 2002[24]	491	8 weeks	Citalopram 40 Escitalopram 20	45.6 vs. 51.2 *P* = NR (ns)	NR	Fair
		8 weeks	Citalopram 40 Escitalopram 10	45.6 vs. 50 *P* = NR (ns)	NR	
Colonna et al., 2005[25]	357	8 weeks	Citalopram 20 Escitalopram 10	55 vs. 63 *P* < 0.05	45 vs. 55 *P* = NR	Fair
		24 weeks	Citalopram 20 Escitalopram 10	78 vs. 80 *P* = NR (ns)	71 vs. 76 *P* = NR	
Lepola et al., 2003[26]	471	8 weeks	Citalopram 20-40 Escitalopram 10-20	52.6 vs. 63.7 *P* = 0.021	42.8 vs. 52.1 *P* = 0.036	Fair
Moore et al., 2005[27]	280	8 weeks	Citalopram 40 Escitalopram 20	61.3 vs. 76.1 *P* = 0.008	43.6 vs. 56.1 *P* = 0.04	Fair
Patris et al., 1996[28]	357	8 weeks	Citalopram 20 Fluoxetine 20	78 vs. 76 *P* = NR (ns)	75 vs. 68 *P* = 0.26	Fair
Haffmans et al., 1996[29]	217	6 weeks	Citalopram 20-40 Fluvoxamine 100-200	30.5 vs. 28.4 *P* = NR	14 vs. 8 *P* = NR (ns)	Fair
Ekselius et al., 1997[30]	400	24 weeks	Citalopram 20-60 Sertraline 50-150	81 vs. 75.5 *P* = NR (ns)	NR	Good
Kasper et al., 2005[31]	518	8 weeks	Escitalopram 10 Fluoxetine 20	46 vs. 37 *P* = NR (ns)	40 vs. 30 *P* = NR (ns)	Fair
Dalery and Honig, 2003[32]	184	6 weeks	Fluoxetine 20 Fluvoxamine 100	60 vs. 60 *P* = NR (ns)	NR	Fair
Rapaport et al., 1996[33]	100	7 weeks	Fluoxetine 20-80 Fluvoxamine 100-150	NR	NR	Fair
Cassano et al., 2002[34]	242	52 weeks	Fluoxetine 20-60 Paroxetine 20-40	NR	NR	Fair
Chouinard et al., 1999[35]	203	12 weeks	Fluoxetine 20-80 Paroxetine 20-50	68.4 vs. 67 *P* = 0.93	59.2 vs. 58 *P* = 0.84	Fair
De Wilde et al., 1993[36]	100	6 weeks	Fluoxetine 20-60 Paroxetine 20-40	62 vs. 67 *P* = NR	NR	Fair
Fava et al., 1998[37]	128	12 weeks	Fluoxetine 20-80 Paroxetine 20-50	NR	NR	Fair
Gagiano et al., 1993[38]	90	6 weeks	Fluoxetine 20-60 Paroxetine 20-40	63 vs. 70 *P* = NR	NR	Fair
Schöne and Ludwig, 1993[39]	108	6 weeks	Fluoxetine 20-60 Paroxetine 20-40	37.5 vs. 16 *P* = 0.03	NR	Fair
Tignol, 1993[40]	178	6 weeks	Fluoxetine 20 Paroxetine 20	78 vs. 75 *P* = NR (ns)	NR	Fair
Fava et al., 2002[41]	284	10-16 weeks	Fluoxetine 20-60 Paroxetine 20-60 Sertraline 50-200	64.8 vs. 68.8 vs. 72.9 *P* = NR	54.4 vs. 57.0 vs. 59.4 *P* = NR	Fair

NR, not reported; ns, not significant.

Table 8. Study characteristics, response and remission rates, and quality ratings of studies in adults with major depressive disorder (continued)

Study	N	Duration	Comparison and Dose (mg/d)	Response (%) and Significance Level	Remission (%) and Significance Level	Quality Rating
Bennie et al., 1995[42]	286	6 weeks	Fluoxetine 20-40 Sertraline 50-100	51 vs. 59 *P* = NR	NR	Fair
Boyer et al., 1998[43]	242	≈ 26 weeks	Fluoxetine 50-150 Sertraline 20-60	42.6 vs. 47.4 *P* = NR	NR	Fair
Newhouse et al., 2000[44,45]	236	12 weeks	Fluoxetine 20-40 Sertraline 50-100	71 vs. 73 *P* = NR	46 vs. 45 *P* = NR	Fair
Sechter et al., 1999[46]	234	24 weeks	Fluoxetine 20-60 Sertraline 50-150	64 vs. 74 *P* = NR	NR	Fair
Van Moffaert et al., 1995[47]	165	8 weeks	Fluoxetine 20 Sertraline 50	Data NR *P* = NR (ns)	NR	Fair
Fava et al., 2000[48]	284	26-32 weeks	Fluoxetine 20-60 Sertraline 50-200 Paroxetine 20-60	NR	NR	Fair
Kroenke et al., 2001[49]	601	36 weeks	Fluoxetine 20 Sertraline 50 Paroxetine 20	NR	NR	Fair
Kiev and Fieger, 1997[50]	60	7 weeks	Fluvoxamine 50-150 Paroxetine 20-50	Data NR *P* = NR (ns)	NR	Fair
Nemeroff et al., 1995[51]	95	7 weeks	Fluvoxamine 50-150 Sertraline 50-200	Data NR *P* = NR (ns)	NR	Fair
Rossini et al., 2005[52]	93	7 weeks	Fluvoxamine 150 Sertraline 200	Data NR *P* = NR (ns)	NR	Fair
Aberg-Wistedt et al., 2000[53]	353	8 weeks	Paroxetine 20-40 Sertraline 50-150	63 vs. 63 *P* = NR (ns)	57.3 vs. 51.6 *P* = NR (ns)	Fair
	353	24 weeks	Paroxetine 20-40 Sertraline 50-150	69 vs. 72 *P* = NR (ns)	73.7 vs. 80.2 *P* = NR (ns)	
SSRIs vs. SSNRIs and SNRIs						
Leinonen et al., 1999[54]	270	8 weeks	Citalopram 20-60 Mirtazapine 15-60	88 vs. 85 *P* = 0.54 Faster onset of mirtazapine	NR	Fair
Allard et al., 2004[55]	150	22 weeks	Citalopram 10-30 Venlafaxine XR 75-150	93 vs. 93 *P* = NR (ns)	23 vs. 19 *P* = NR (ns)	Fair
Bielski et al., 2004[56]	198	8 weeks	Escitalopram 20 Venlafaxine XR 225	61 vs. 48 *P* = NR (ns)	36 vs. 32 *P* = NR (ns)	Fair
Montgomery et al., 2004[57]	293	8 weeks	Escitalopram 10-20 Venlafaxine XR 75-150	77 vs. 80 *P* = NR (ns)	70 vs. 70 *P* = NR (ns)	Fair
Goldstein et al., 2002[58]	173	8 weeks	Fluoxetine 20 Duloxetine 40-120	45 vs. 49 *P* = 0.39	30 vs. 43 *P* = 0.82	Fair
Hong et al., 2003[59]	133	6 weeks	Fluoxetine 20-40 Mirtazapine 15-45	51 vs. 58 *P* = NR (ns) Faster onset of mirtazapine	27 vs. 35 *P* = NR (ns)	Fair
Versiani et al., 2005[60]	297	8 weeks	Fluoxetine 20-40 Mirtazapine 15-60	Data NR *P* = NR (ns) Faster onset of mirtazapine	41.4 vs. 40.1 *P* = NR (ns)	Fair

Table 8. Study characteristics, response and remission rates, and quality ratings of studies in adults with major depressive disorder (continued)

Study	N	Duration	Comparison and Dose (mg/d)	Response (%) and Significance Level	Remission (%) and Significance Level	Quality Rating
Wheatley et al., 1998[61]	133	6 weeks	Fluoxetine 20-40 Mirtazapine 15-60	Data NR *P* = NR (ns) Faster onset of mirtazapine	25.4 vs. 23.3 *P* = NR (ns)	Fair
Alves et al., 1999[62]	87	12 weeks	Fluoxetine 20-40 Venlafaxine 75-150	74 vs. 87 *P* = NR	41 vs. 51 *P* = NR	Fair
Costa e Silva, 1998[63]	382	8 weeks	Fluoxetine 20-40 Venlafaxine 75-225	82 vs. 86.8 *P* = 0.074	60.2 vs. 60.2 *P* = NR	Fair
De Nayer et al., 2002[64]	146	12 weeks	Fluoxetine 20-40 Venlafaxine 75-150	49.3 vs. 75 *P* = 0.001	40.3 vs. 59.4 *P* = 0.028	Fair
Dierick et al., 1996[65]	314	8 weeks	Fluoxetine 20 Venlafaxine 75-150	60 vs. 72 *P* = 0.023 (at week 6)	NR	Fair
Nemeroff and Thase, 2005[66]	308	6 weeks	Fluoxetine 20-60 Venlafaxine 75-225	45 vs. 53 *P* = 0.034	28 vs. 32 *P* = 0.250	Fair
Rudolph and Feiger, 1999[67]	301	8 weeks	Fluoxetine 20-60 Venlafaxine XR 75-225	50 vs. 57 *P* = 0.07	22 vs. 37 *P* < 0.05	Fair
Silverstone and Ravindran, 1999[68]	368	12 weeks	Fluoxetine 20-60 Venlafaxine XR 75-225	62 vs. 67 *P* < 0.05	NR	Fair
Tzanakaki et al., 2000[69]	109	6 weeks	Fluoxetine 60 Venlafaxine 225	66 vs. 70 *P* = NR	36 vs. 41 *P* = NR	Fair
Tylee et al., 1997[70]	341	12 weeks	Fluoxetine 20 Venlafaxine 75	62.8 vs. 55.1 *P* = NR	34.1 vs. 35.4 *P* = NR (ns)	Fair
Detke et al., 2004[71]	367	8 weeks	Paroxetine 20 Duloxetine 80 Duloxetine 120	74 vs. 65 vs. 71 *P* = NR (ns)	44 vs. 46 vs. 52 *P* = NR (ns)	Fair
Benkert et al., 2000[72]	275	6 weeks	Paroxetine 20-40 Mirtazapine 15-45	53.7 vs. 58.3 *P* = NR (ns) Faster onset of mirtazapine	34.1 vs. 40.9 *P* = NR (ns)	Fair
Schatzberg et al., 2002[73]	255	8 weeks	Paroxetine 20-40 Mirtazapine 15-45	56.7 vs. 64.0 *P* = NR (ns) Faster onset of mirtazapine	NR	Fair
Ballus et al., 2000[74]	84	12 weeks	Paroxetine 20-40 Venlafaxine 75-150	NR *P* = NR (ns)	33 vs. 57 *P* = 0.011	Fair
		24 weeks	Paroxetine 20-40 Venlafaxine 75-150	49 vs. 59 *P* = NR (ns)	NR *P* = NR (ns)	
McPartlin et al., 1998[75]	361	12 weeks	Paroxetine 20 Venlafaxine XR 75	76 vs. 76 *P* = NR (ns)	46 vs. 48 *P* = NR (ns)	Fair
Behnke et al., 2003[76]	345	8 weeks	Sertraline 50-150 Mirtazapine 30-45	NR *P* = NR (ns) Faster onset of mirtazapine	NR	Fair
Mehtonen et al., 2000[77]	147	8 weeks	Sertraline 50-100 Venlafaxine 75-150	68 vs. 83 *P* = 0.05	45 vs. 68 *P* = 0.008	Good
Sir et al., 2005[78]	163	8 weeks	Sertraline 50-150 Venlafaxine XR 75-225	70.9 vs. 70.9 *P* = 0.95	59.5 vs. 54.4 *P* = 0.47	Good

Table 8. Study characteristics, response and remission rates, and quality ratings of studies in adults with major depressive disorder (continued)

Study	N	Duration	Comparison and Dose (mg/d)	Response (%) and Significance Level	Remission (%) and Significance Level	Quality Rating

SSRIs vs. other second-generation antidepressants						
Coleman et al., 2001[79]	456	8 weeks	Fluoxetine 20-60 Bupropion SR 150-400	57 vs. 56 *P* = NR (ns)	40 vs. 47 *P* = NR (ns)	Fair
Feighner et al., 1991[80]	123	6 weeks	Fluoxetine 20-80 Bupropion 225-450	58 vs. 63 *P* = NR (ns)	NR	Fair
Rush et al., 1998[81]	85	8 weeks	Fluoxetine 20-40 Nefazodone 200-500	45 vs. 47 *P* = NR (ns)	NR	Fair
Beasley et al., 1991[82]	126	6 weeks	Fluoxetine 20-60 Trazodone 100-400	62 vs. 69 *P* = NR (ns)	51 vs. 42 *P* = NR (ns)	Fair
Perry et al., 1989[83]	40	6 weeks	Fluoxetine 20-60 Trazodone 50-400	NR	NR	Fair
Weihs et al., 2000[84]	100	6 weeks	Paroxetine 10-40 Bupropion SR 100-300	77 vs. 71 *P* = NR (ns)	NR	Fair
Baldwin et al., 1996[85]	206	8 weeks	Paroxetine 20-40 Nefazodone 200-600	60 vs. 58 *P* = NR (ns)	NR	Fair
Hicks et al., 2002[86]	40	8 weeks	Paroxetine 20-40 Nefazodone 400-600	*P* = NR (ns)	NR	Fair
Kasper et al., 2005[87]	108	6 weeks	Paroxetine 20-40 Trazodone 150-450	91 vs. 87 *P* = NR (ns)	68 vs. 69 *P* = NR (ns)	Fair
Coleman et al., 1999[88]	364	8 weeks	Sertraline 50-200 Bupropion SR 150-400	61 vs. 66 *P* = NR (ns)	NR	Fair
Croft et al., 1999[89]	360	8 weeks	Sertraline 50-200 Bupropion SR 150-400	68 vs. 66 *P* = NR (ns)	NR	Fair
Kavoussi et al., 1997[90] Rush et al., 2001[91]	248	16 weeks	Sertraline 50-200 Bupropion SR 100-300	NR	NR	Fair
Feiger et al., 1996[92]	160	6 weeks	Sertraline 50-200 Nefazodone 100-600	57 vs. 59 *P* = NR (ns)	NR	Fair
SSNRIs and SNRIs vs. SNRIs						
Guelfi et al., 2001[93]	157	8 weeks	Mirtazapine 45-60 Venlafaxine 225-375	62 vs. 52 *P* = NR (ns)	NR	Fair
SNRIs vs. other second-generation antidepressants						
Halikas, 1995[94]	150	6 weeks	Mirtazapine 5-35 Trazodone 40-280	51 vs. 41 *P* = NR (ns)	NR	Fair
van Moffaert et al., 1995[95]	200	6 weeks	Mirtazapine 24-72 Trazodone 150-450	61 vs. 51 *P* = NR	NR	Fair
Cunningham et al., 1994[96]	225	6 weeks	Venlafaxine 75-200 Trazodone 150-400	72 vs. 60 *P* = NR (ns)	NR	Fair
Other second-generation antidepressants vs. other second-generation antidepressants						
Weisler et al., 1994[97]	124	6 weeks	Bupropion 225-450 Trazodone 150-400	55.9 vs. 40.4 *P* = NR	46 vs. 31 *P* = NR	Fair

Note: Venlafaxine not otherwise specified refers to the immediate-release formulation (given twice a day), venlafaxine XR is the extended-release formulation.

Table 9. Possible comparisons of second-generation antidepressants involving SSRIs and number of included head-to-head trials

Comparison	Number of Studies	Comparison	Number of Studies
SSRIs vs. SSRIs		**SSRIs vs. Other Second-Generation Antidepressants**	
Citalopram vs. Escitalopram	4	Citalopram vs. Bupropion	0
Citalopram vs. Fluoxetine	1	Citalopram vs. Nefazodone	0
Citalopram vs. Fluvoxamine	1	Citalopram vs. Trazodone	0
Citalopram vs. Paroxetine	0	Escitalopram vs. Bupropion	0
Citalopram vs. Sertraline	1	Escitalopram vs. Nefazodone	0
Escitalopram vs. Fluoxetine	1	Escitalopram vs. Trazodone	0
Escitalopram vs. Fluvoxamine	0	*Fluoxetine vs. Bupropion*	2
Escitalopram vs. Paroxetine	0	*Fluoxetine vs. Nefazodone*	1
Escitalopram vs. Sertraline	0	*Fluoxetine vs. Trazodone*	2
Fluoxetine vs. Fluvoxamine	2	Fluvoxamine vs. Bupropion	0
Fluoxetine vs. Paroxetine	10	Fluvoxamine vs. Nefazodone	0
Fluoxetine vs. Sertraline	8	Fluvoxamine vs. Trazodone	0
Fluvoxamine vs. Paroxetine	1	*Paroxetine vs. Bupropion*	1
Fluvoxamine vs. Sertraline	2	*Paroxetine vs. Nefazodone*	2
Paroxetine vs. Sertraline	4	*Paroxetine vs. Trazodone*	1
SSRIs vs. SSNRIs		*Sertraline vs. Bupropion*	3
Citalopram vs. Duloxetine	0	*Sertraline vs. Nefazodone*	1
Escitalopram vs. Duloxetine	0	Sertraline vs. Trazodone	0
Fluoxetine vs. Duloxetine	1		
Fluvoxamine vs. Duloxetine	0		
Paroxetine vs. Duloxetine	1		
Sertraline vs. Duloxetine	0		
SSRIs vs. SNRIs			
Citalopram vs. Mirtazapine	1		
Citalopram vs. Venlafaxine	1		
Escitalopram vs. Mirtazapine	0		
Escitalopram vs. Venlafaxine	2		
Fluoxetine vs. Mirtazapine	3		
Fluoxetine vs. Venlafaxine	9		
Fluvoxamine vs. Mirtazapine	0		
Fluvoxamine vs. Venlafaxine	0		
Paroxetine vs. Mirtazapine	2		
Paroxetine vs. Venlafaxine	2		
Sertraline vs. Mirtazapine	1		
Sertraline vs. Venlafaxine	2		

Table 10. Possible comparisons of second-generation antidepressants involving SSNRIs, SNRIs, and other antidepressants and number of included head-to-head trials

Comparison	Number of Studies
SSNRIs and SNRIs vs. SNRIs	
Duloxetine vs. Venlafaxine	0
Duloxetine vs. Mirtazapine	0
Mirtazapine vs. Venlafaxine	1
SSNRIs vs. Other Second-Generation Antidepressants	
Duloxetine vs. Bupropion	0
Duloxetine vs. Nefazadone	0
Duloxetine vs. Trazodone	0
SNRIs vs. Other Second-Generation Antidepressants	
Mirtazapine vs. Bupropion	0
Mirtazapine vs. Nefazodone	0
Mirtazapine vs. Trazodone	2
Venlafaxine vs. Bupropion	0
Venlafaxine vs. Nefazadone	0
Venlafaxine vs. Trazodone	1
Other Second-Generation Antidepressants vs. Other Second-Generation Antidepressants	
Bupropion vs. Nefazadone	0
Bupropion vs. Trazodone	1
Nefazadone vs. Trazodone	0

ITT analysis. Few authors, however, reported the overall number of patients lost to followup from randomization to the end of the trial. In addition, many studies did not report the ethnic backgrounds of participants.

Loss to followup (number of patients randomized who did not proceed to endpoint), a potential source of bias, was a frequent problem for internal validity. Only 14 trials (17.5 percent) reported a loss to followup of less than 20 percent. The high drop-out rates for many studies may be attributable to specific characteristics of a psychiatric outpatient population and a high rate of adverse events in the examined drug class.

Major Depressive Disorders: Key Points

Seventy-two head-to-head comparisons (Table 9) were available for a total of 35 potential comparisons between the 12 second-generation antidepressants addressed in this report. Of these, only five trials[93-97] directly compared any non-SSRI second-generation antidepressant to any other non-SSRI agent (Table 10); of these, only one comparison was evaluated in more than one trial. The strength of evidence, overall for comparative efficacy and effectiveness, was rated moderate. Overall, 38 percent of patients did not achieve a treatment response during 6 weeks to 12 weeks of treatment with second-generation antidepressants; 54 percent did not achieve remission.

Direct evidence was considered sufficient to conduct meta-analyses for four drug-drug comparisons:

- Citalopram vs. escitalopram (four published studies[24-27] and one FDA review;[98] 1,545 patients): Patients on escitalopram had an additional treatment effect of a 1.25 point reduction (95% CI, 0.10-2.39; *P* = 0.02) on the MADRS compared with patients on citalopram. The relative risk of response was statistically significantly greater for escitalopram than for citalopram (RR, 1.14; 95% CI, 1.04-1.26). The number needed to treat (NNT) to gain one additional responder at week 8 with escitalopram was 14 (95% CI, 7-111). Both drugs are produced by the same manufacturer, which funded all available studies.

- Fluoxetine vs. paroxetine (seven studies;[35-41] 950 patients): We did not find any statistically significant differences in either effect sizes on HAM-D or response rates between fluoxetine and paroxetine. Fluoxetine had an additional reduction of 0.55 (95% CI, -1.4-0.36; *P* = 0.23) points on HAM-D compared with paroxetine; paroxetine led to a higher rate of responders than fluoxetine (RR, 1.09; 95% CI, 0.99-1.21).

- Fluoxetine vs. sertraline (four studies;[41,42,44,46] 940 patients): Patients on sertraline had an additional, statistically nonsignificant treatment effect of a 0.75 point reduction (95% CI, -0.45-1.95) on the HAM-D scale compared with patients on fluoxetine. The relative risk of response was statistically significantly greater for sertraline than for fluoxetine (RR, 1.11; 95% CI, 1.01-1.21). The NNT to gain one additional responder at 6 to 12 weeks with sertraline was 14 (95% CI, 8-22).

- Fluoxetine vs. venlafaxine (eight studies;[62-67,69,70] 1,814 patients): Patients on venlafaxine had an additional, statistically nonsignificant treatment effect of a 1.31 point reduction (95% CI, 0.10-2.39; *P* = 0.13) on the HAM-D scale compared with patients on fluoxetine. The relative risk of response was statistically significantly greater for venlafaxine than for fluoxetine (RR, 1.12; 95% CI, 1.01-1.24). The NNT to gain one additional responder at 6 to 12 weeks with venlafaxine was 12 (95% CI, 7-50). All studies were funded by the makers of venlafaxine.

Very few comparative effectiveness trials were available; their findings were generally consistent with efficacy trials.[30,46,49]

Findings from indirect comparisons yielded no statistically significant differences in response rates among other potential comparisons. Although the precision of some estimates was low, leading to inconclusive results, treatment effects are similar across all comparisons.

Eighteen studies (N = 4,050) comparing one second-generation antidepressant with another indicated no differences in health-related quality of life.[24,43,45,46,53,54,56,60,61,66,75,78,82,84,93,99-101] Quality of life, however, was rarely assessed as a primary outcome measure. The strength of evidence is moderate.

Seven studies, all funded by the maker of mirtazapine, reported that mirtazapine has a significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline (Table 11).[54,59-61,72,73,76] The NNT to yield one additional responder after 1 or 2 weeks of treatment is 7 (95% CI, 5-12). This treatment effect was consistent across all studies. The strength of evidence is moderate.

Five trials provide evidence that bupropion leads to greater satisfaction with sexual activity than sertraline[88-90] and fluoxetine (Table 12).[79,80] The NNT to yield one additional person with a

high overall satisfaction of sexual functioning is 7. This treatment effect was consistent across all studies.

We did not find any efficacy evidence addressing Key Question 1b.

Table 11. Characteristics of trials comparing mirtazapine to SSRIs on onset of action (response rate)

Study	Sample Size	Comparison	Effect Size	*P*-value	Comments
Leinonen et al., 1999[54]	270	Citalopram	Significantly greater reduction of HAM-D scores with mirtazapine at day 14 (difference: -2.3)	*P* = 0.002	No statistically significant differences in response and remission rates at endpoint
Hong et al., 2003[59]	133	Fluoxetine	At day 28 significantly more responders with mirtazapine (53.3% vs. 39.0%) RRR, 0.23 RD: 0.14 NNT: 7	*P* = NR (ns)	No statistically significant differences in overall response rate at week 6; more responders in the mirtazapine group (58% vs. 51%)
Versiani et al., 2005[60]	297	Fluoxetine	Significantly more responders at day 7 with mirtazapine (data NR) Higher rate of remitters for mirtazapine at days 14 (6.2 % vs. 2.0%), 28 (18.6% vs. 12.9%), and 42 (29.0% vs. 21.1%)	*P* = 0.002 *P* = NR (ns)	No statistically significant differences in response and remission at endpoint (day 42)
Wheatley et al., 1998[61]	133	Fluoxetine	Significantly more responders at day 28 with mirtazapine (data NR)	*P* = 0.006	Statistically significantly greater decrease of HAM-D scores for mirtazapine at days 21 and 28. No statistically significant differences in response and remission at endpoint (day 56)
Benkert et al., 2000[72]	275	Paroxetine	Significantly more responders (23.2% vs. 8.9%) and remitters (8.8% vs. 2.4%) at day 7 with mirtazapine. RRR, 0.15 0.07 RD: 0.14 0.07 NNT: 8 15	Response: *P* = 0.002 Remission: *P* = 0.03	More responders and remitters in the mirtazapine group throughout the study. No statistically significant difference at endpoint (response: 58.3% vs. 53.7%; remission: 40.9% vs. 34.8%)
Schatzberg et al., 2002[73]	255	Paroxetine	Significantly more responders at day 14 with mirtazapine (27.8% vs. 13.3%) RRR, 0.17 RD: 0.14 NNT: 7 Significantly greater decrease of HAM-D scores from day 7 to day 21 with mirtazapine Median time to response: Mirtazapine: 26 days Paroxetine: 40 days	*P* = 0.005 *P* < 0.01 (day 7, 14) *P* = 0.024 (day 21) Kaplan-Mayer: *P* = 0.016	No statistically significant differences in overall response rate at week 8; more responders in the mirtazapine group (51% vs. 8%) at endpoint

Behnke et al., 2003[76]	346	Sertraline	Significantly higher response rates at days 7, 10, and 14 with mirtazapine (data NR)	day 7: $P < 0.05$ day 10: $P < 0.01$ day 14: $P < 0.05$	No statistically significant differences in response and remission at endpoint (day 56)

HAM-D, Hamilton Rating Scale for Depression; NNT, number needed to treat; NR, not reported; ns, not significant; RD, risk difference; RRR, relative risk reduction.

Table 12. Characteristics of trials comparing bupropion to SSRIs on sexual functioning and satisfaction

Study	Sample Size	Comparison	Effect Size	*P*-value	Comments
Feighner et al., 1991[80]	61	Fluoxetine	Higher rates for fluoxetine of impotence (4.7% vs. 0%) anorgasmia (1.7% vs. 0%) libido decrease (1.7% vs. 0%)	NR	Self-reporting of sexual adverse events
Coleman et al., 2001[79]	456	Fluoxetine, placebo	Significantly more bupropion SR patients were satisfied with overall sexual functioning (analysis only for patients satisfied at baseline; no data reported)	$P < 0.05$	DSM-IV criteria for sexual dysfunction disorders No statistically significant differences in efficacy outcome measures at endpoint (week 8)
Coleman et al., 1999[88]	364	Sertraline	Beginning at day 21, significantly more patients on bupropion SR were satisfied with their sexual functioning (endpoint: 85% vs. 62%) Endpoint: RRR, 0.59 RD: 0.22 NNT: 5	$P < 0.05$	DSM-IV criteria for sexual dysfunction disorders No statistically significant differences in efficacy outcome measures at endpoint (week 8)
Croft et al., 1999[89]	360	Sertraline, placebo	Beginning at day 7 through day 42 significantly more bupropion SR patients were satisfied with overall sexual functioning; difference was not statistically significant at endpoint (75% vs. 65%) endpoint: RRR, 0.29 RD: 0.10 NNT: 10	$P < 0.05$	Sexual function assessed in investigator-conducted structured interview No statistically significant differences in efficacy outcome measures at endpoint (week 8)
Kavoussi et al., 1997[90,102]	248	Sertraline	Significantly more patients on sertraline experienced orgasm delays and/or failure Women: 41% vs. 7% RRR, 0.85 RD: 0.38 NNT: 3 Men: 61% vs. 10% RRR, 0.84 RD: 0.51 NNT: 2 Higher overall satisfaction with sexual functioning with bupropion SR at endpoint (79% vs. 58%) RRR, 0.50 RD: 0.21 NNT: 5	$P < 0.01$ $P < 0.001$	Sexual function assessed in investigator-conducted structured interview No statistically significant differences in efficacy outcome measures at endpoint (week 16)

DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; NNT, number needed to treat; NR, not reported; RD, risk difference; RRR, relative risk reduction.

Major Depressive Disorder: Detailed Analysis

Head-to-head evidence: SSRIs vs. SSRIs. *Citalopram vs. escitalopram.* Four published trials[24-27] and one unpublished[98] trial compared the efficacy of citalopram and escitalopram. Four studies were conducted over 8 weeks[24,26,27,98] and one over 24 weeks.[25] One study was a flexible dose trial.[26] Table 13 summarizes study characteristics and differences in effect sizes of studies comparing citalopram with escitalopram.

Table 13. Characteristics and effect sizes of studies comparing citalopram with escitalopram

Study	N	Duration	Dosage Cit. - Esc. mg/d	Response (%)	Remission (%)	Quality Rating
Burke et al., 2002[24]	491	8 weeks	40 vs. 20	45.6 vs. 51.2 *P* = NR (ns)	NR	Fair
			40 vs. 10	45.6 vs. 50 *P* = NR (ns)	NR	
Colonna et al., 2005[25]	357	8 weeks	20 vs. 10	55 vs. 63 *P* < 0.05	NR	Fair
		24 weeks	20 vs. 10	78 vs. 80 *P* = NR (ns)	NR	
Lepola et al., 2003[26]	471	8 weeks	20-40 vs. 10-20	52.6 vs. 63.7 *P* = 0.021	42.8 vs. 52.1 *P* = 0.036	Fair
Moore et al., 2005[27]	280	8 weeks	40 vs. 20	61.5 vs. 76.1 *P* = 0.009	43.6 vs. 56.1 *P* = 0.04	Fair
Unpublished Study SCT MD-02[98]	375	8 weeks	20-40 vs. 10-20	51 vs. 46 *P* = NR	NR	Fair

NR, not reported; ns, not significant.

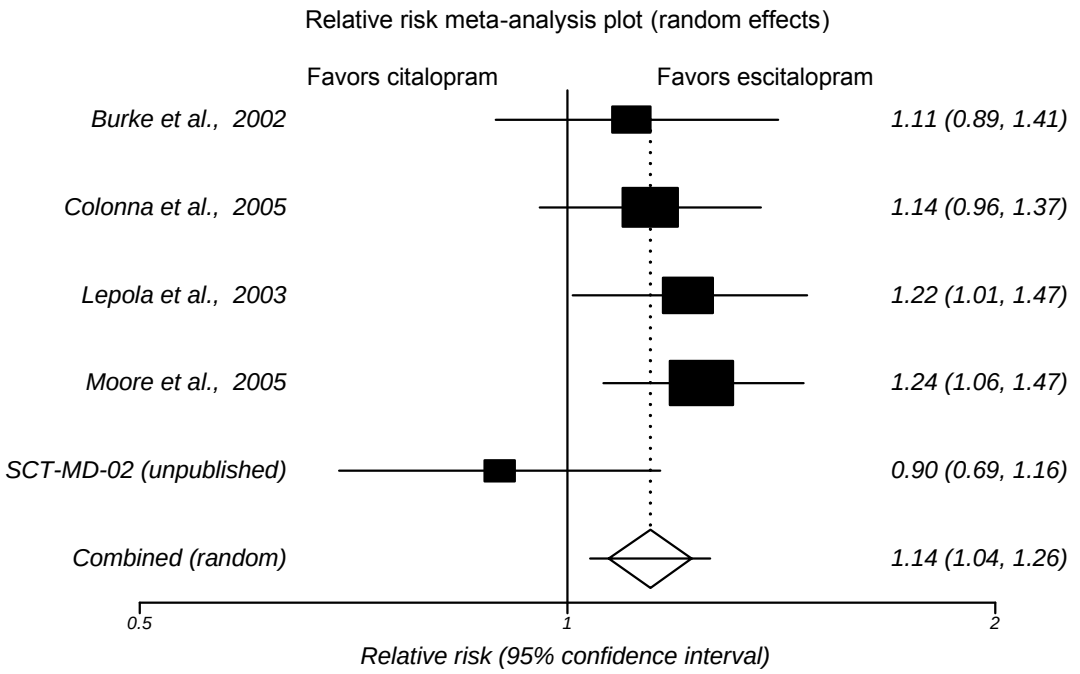
Overall, results of individual studies favored escitalopram over citalopram. In three studies, differences in response rates reached statistical significance at 8 weeks.[25-27] The flexible dose trial was a European-Canadian study that compared efficacy and harms of citalopram (20-40 mg/day), escitalopram (10-20 mg/day), and placebo in 471 depressed outpatients attending primary care centers.[26] ITT results showed that the escitalopram group had significantly more patients responding (63.7 percent vs. 52.6 percent; $P = 0.021$) and achieving remission (52.1 percent vs. 42.8 percent; $P < 0.036$) than the citalopram group. Escitalopram was numerically better at all time points on three scales (MADRS, Clinical Global Impressions Improvement Scale [CGI-I], Clinical Global Impressions Severity Scale [CGI-S]). The study did not assess health outcomes.

The 24-week study was a fixed-dose trial (escitalopram 10 mg/day, citalopram 20 mg/day) of 357 European primary care patients over 24 weeks.[25] Escitalopram patients had significantly higher response rates at week 8 (63 percent vs. 55 percent; $P < 0.05$) but not at week 24 (80 percent vs. 78 percent; *P* = NR). Escitalopram had significantly lower CGI-S scores (1.75 vs. 2.00) and significantly fewer withdrawals (12.7 percent vs. 22.4 percent) than citalopram at week 24.

We conducted two meta-analyses of these studies comparing the effects of citalopram with those of escitalopram on MADRS scores at week 8. The outcome of the first meta-analysis was the relative risk of being a responder on the MADRS scale at week 8 (Figure 3). In addition to

the four published trials, we included data from one unpublished study from the FDA Center for Drug Evaluation and Research (CDER) database.[98] A "response" was defined as an improvement of 50 percent or more on the MADRS. Pooled results included 1,545 patients and yielded a statistically significant additional treatment effect for escitalopram. The relative risk that a patient would respond was 1.14 (95% CI, 1.04-1.26) for escitalopram relative to citalopram. Both random effects and fixed effects models presented similar, statistically significant results. The NNT to gain one additional responder based on the pooled risk difference is 14 (95% CI, 7-111).

Figure 3. Relative risk meta-analysis of MADRS response rates comparing citalopram with escitalopram



The second meta-analysis was an effect size meta-analysis of all five studies assessing the pooled difference of points on the MADRS (Figure 4). Overall, this analysis included data on 1,545 patients. The weighted mean difference (WMD) presented an additional treatment effect of a 1.13 point reduction (95% CI, 0.18-2.09; $P = 0.02$) for escitalopram compared with citalopram. Although the difference was statistically significant, the clinical implications remain to be clarified. A 1.13 point change on the MADRS represents about one-fifth to one-quarter of a standard deviation. A recent methods study concluded that, in general, a change of about one-half of a standard deviation on a health-related scale reflects a minimally important difference for a patient.[103]

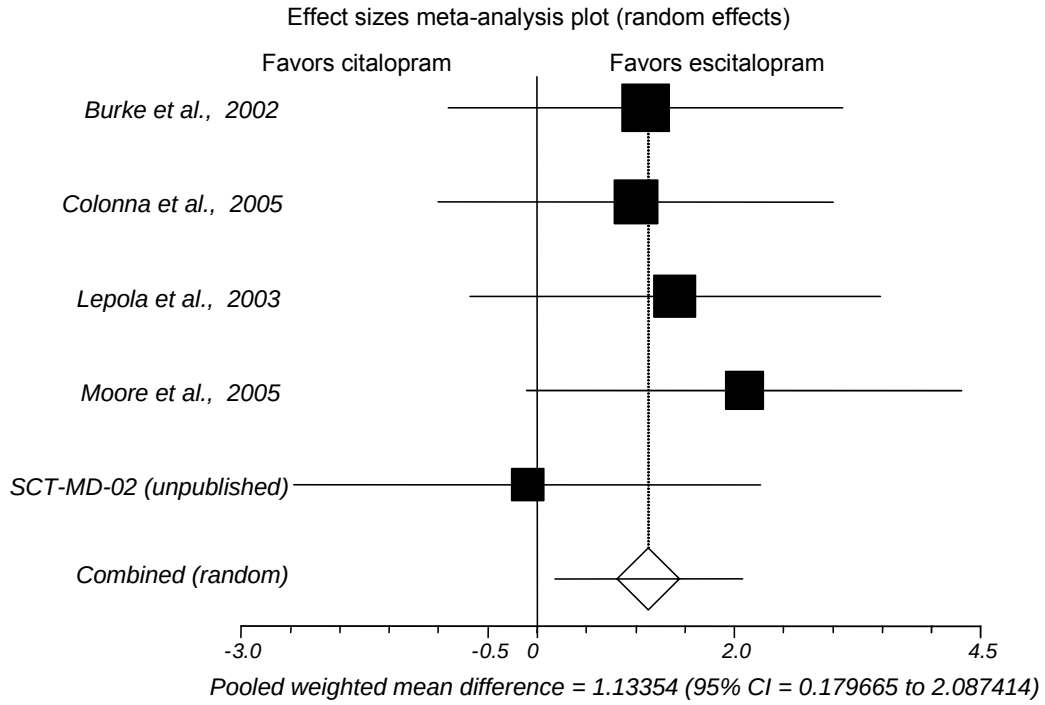
Both citalopram and escitalopram are produced by the same manufacturer, which funded all available studies. Generic brands of citalopram are available in the United States; escitalopram is still under patent protection.

Citalopram vs. fluoxetine. In a French trial, 397 outpatients with MDD attending general practices were randomly assigned to citalopram (20 mg/day) or fluoxetine (20 mg/day) over 8

weeks.[28] Citalopram had a faster onset of efficacy than fluoxetine; significantly more patients were rated as responding (35 percent vs. 24 percent; $P = 0.048$) or completely recovered (27 percent vs. 16 percent; $P = 0.034$) on the MADRS after 2 weeks. At 8 weeks, however, response rates for the citalopram and the fluoxetine group were similar (78 percent vs. 76 percent; P = NR).

Citalopram vs. fluvoxamine. A Dutch study (N = 217) did not find any differences in efficacy (HAM-D, CGI, Zung self-rating depression scale at 6 weeks) between citalopram (20-40 mg/day) and fluvoxamine (100-200 mg/day).[29] Remission rates did not differ significantly between citalopram and fluvoxamine treatments (14 percent vs. 8 percent; P = NR).

Figure 4. Effect size meta-analysis comparing citalopram with escitalopram on the MADRS



Citalopram vs. sertraline. A Swedish study rated good quality assessed the effectiveness of citalopram (20-60 mg/day) and sertraline (50-150 mg/day) in 400 patients in general practice during 24 weeks of treatment.[30] The majority of patients suffered recurrent depression (sertraline, 56 percent; citalopram, 65 percent) and used other medications for medical illnesses (sertraline, 55 percent; citalopram, 44.5 percent). The investigators found no significant differences between treatment groups in any outcome measures at any point in time (MADRS, CGI-S, CGI-I). Also, in a subgroup analysis of patients with recurrent depression and single episode depression, they did not report any differences in effectiveness between drugs. Response rates were similar at week 24 (citalopram, 81.0 percent; sertraline, 75.5 percent; P = NR). This study was one of only a few trials not funded by the pharmaceutical industry; it can be considered an effectiveness trial.

Escitalopram vs. fluoxetine. A multinational RCT enrolled patients older than 65 years (n = 518) in general-practice and psychiatric-specialist settings to assess the comparative efficacy of escitalopram (10 mg/day) and fluoxetine (20 mg/day).[31] Both treatment groups had no greater efficacy than the placebo control group after 8 weeks of treatment. Response and remission rates did not differ significantly between the active treatment groups.

Fluoxetine vs. fluvoxamine. Two studies evaluated the comparative efficacy and safety of fluoxetine and fluvoxamine in 284 outpatients with MDD.[32,33] A 7-week flexible-dose study (fluoxetine: 20-80 mg/day; fluvoxamine 100-150 mg/day) did not identify any statistically or clinically significant differences in efficacy between the two treatment groups (HAM-D, HAM-A, CGI-S, Raskin-Covi Scale, Hopkins Symptoms Checklist [HSCL-D20]).[33] Both treatment regimens significantly improved scores on assessment scales over 7 weeks.

In a 6-week fixed-dose European trial (fluoxetine 20 mg/day; fluvoxamine 100 mg/day) in 184 outpatients with MDD,[32] results are consistent with those of the flexible-dose study; scores on the primary outcome measure (HAM-D) were not significantly different at any time. At endpoint, the drugs were equally effective for secondary outcome measures such as suicidal ideation, sleep, anxiety, and severity of illness (CGI, Clinical Anxiety Scale [CAS], the Irritability, Depression, and Anxiety Scale [IDAS], Beck's Scale for Suicide Ideation [Beck's SSI]). Fluvoxamine had significantly more responders on the CGI-S (29 percent vs. 16 percent; $P < 0.05$) and a greater reduction of CGI-S scores ($P < 0.05$) at week 2 but not at weeks 4 or 6.

Fluoxetine vs. paroxetine. Ten studies compared fluoxetine to paroxetine.[34-41,48,49] Most studies lasted from 6 to 12 weeks. Efficacy measures included HAM-D, HAM-A, MADRS, CGI-S, CGI-I, Covi Anxiety Scale, and others. Overall, these studies did not indicate substantial differences in outcome measures between fluoxetine and paroxetine. The largest study was a Canadian RCT (n = 203) with a study duration of 12 weeks.[35] At study endpoint, fluoxetine (20-80 mg/day) and paroxetine (20-50 mg/day) presented similar response (68 percent vs. 67 percent; $P = 0.93$) and remission rates (59 percent vs. 58 percent; $P = 0.84$).

One study was conducted in an inpatient population.[40] Results were consistent with findings of the other studies.

We conducted a meta-analysis of these seven studies (excluding three that did not report data[34,48,49]) using HAM-D scores at the end of followup.[35-41] We defined "response" as an improvement of 50 percent or more on the HAM-D. The statistical analysis included 950 patients. The pooled estimate of the random effects model, presented in Figure 5, indicates that fluoxetine and paroxetine do not differ significantly in efficacy (RR, 1.09; 95% CI, 0.99-1.21). Removing the study conducted in an inpatient population[40] from the analysis did not change the point estimate. An effect size meta-analysis (Figure 6) also did not detect a statistically significant difference between fluoxetine and paroxetine (-0.55; 95% CI, -1.46-0.36).

Figure 5. Relative risk meta-analysis of response rates comparing fluoxetine with paroxetine on the HAM-D

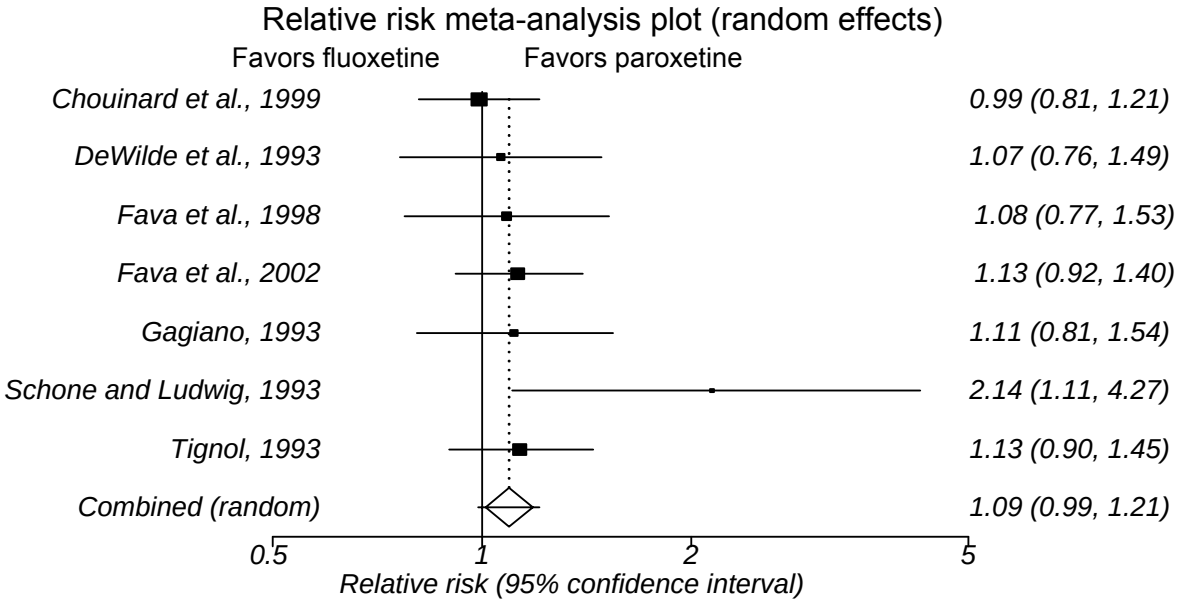
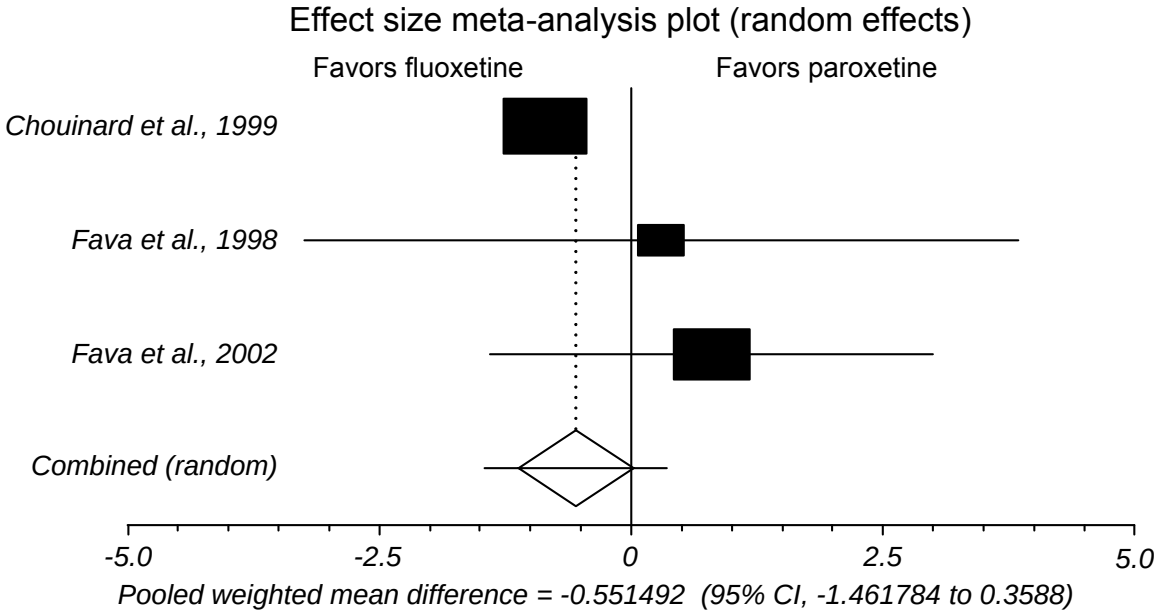


Figure 6. Effect size meta-analysis comparing fluoxetine with paroxetine on the HAM-D



Two RCTs were conducted in a population older then 60 years.[34,39] An Italian study lasting 1 year enrolled 242 patients to compare the effects of fluoxetine (20-60 mg/day) and paroxetine (20-40 mg/day) on depressive symptoms, mood, and cognitive function in nondemented persons (65 years or older).[34]

In both trials, paroxetine-treated patients achieved higher response rates than patients in the fluoxetine group. In one study, differences in response rates reached statistical significance (37.5

percent vs. 17.5 percent; $P = 0.04$).[39] In the long-term Italian study, treatment groups did not differ significantly at study endpoint on CGI scores or most cognitive scales (Blessed Information and Memory Test [BIMT], Mini-Mental State Examination [MMSE], Clifton Assessment Schedule [CLAS]).[34]

Five studies did not detect differences between fluoxetine and paroxetine in improvement of anxiety in patients with depression (HAM-A, Covi Anxiety Scale).[34,35,37,38,41]

Fluoxetine vs. sertraline. Eight studies compared fluoxetine with sertraline.[41-49] The best evidence consisted of two effectiveness trials[46,49] and one efficacy trial[43] with long periods of followup.

Two multicenter trials in France comparing fluoxetine (20-60 mg/day) and sertraline (50-150 mg/day) were conducted in office settings (private psychiatrists and general physicians [GPs]).[43,46] The psychiatrist study[46] randomized 238 patients for 24 weeks; the GP study[43] randomized 242 patients for nearly 26 weeks (180 days). The majority of patients had concomitant medical conditions. Both studies assessed quality of life as a secondary outcome measure (Sickness Impact Profile [SIP], Functional Status Questionnaire [FSQ]). Exclusion criteria were less stringent in the GP trial than the psychiatrist trial. Loss to followup was 4.5 percent in the GP trial and 29.8 percent in the psychiatrist trial. In the GP trial, researchers conducted outcome assessments only at day 120 and day 180, but patients could choose to consult the physician at any time. ITT analyses in both studies did not reveal any statistically significant differences in any primary (MADRS, HAM-D, CGI) or secondary (Covi Anxiety Scale, HAD, SIP, Leeds Sleep Evaluation) efficacy measures or in the incidence of adverse events.

The ARTIST (A Randomized Trial Investigating SSRI Treatment) trial was an open-label RCT designed as an effectiveness study and carried out in primary care physician settings over 9 months.[49] This study did not meet our eligibility criteria because it was an open-label trial; we present it because it is one of only a few effectiveness trials. This study enrolled 601 patients at 76 sites. Initial diagnosis for enrollment was not based on diagnostic criteria but rather on the judgment of the treating physician. Criteria-based evaluation classified 74 percent of patients as having MDD, 18 percent dysthymia, and 8 percent minor depression. Patients' treatments could be switched among study drugs or to other antidepressive medications as needed. ITT analysis maintained the original randomization. Outcome measures assessing changes in depression and health-related quality of life measures (work, social and physical functioning, concentration and memory, and sexual functioning) were administered over the telephone by a blinded third party. Range of dosage and loss to followup were incompletely reported.

Results of the ARTIST trial did not reveal any significant differences among drugs in any outcome measures at either 3 or 9 months.[49] Compared with baseline measures, all treatment groups significantly improved during the study. Subgroup analyses did not show different effectiveness for patients with MDD or for those older than 60 years.

Four additional trials did not find any significant differences in primary outcome measures (HAM-D, MADRS, CGI-S).[41,42,44,45,47] Studies lasted from 6 weeks to 16 weeks. One study was conducted in 236 participants older than 60 years.[44,45] In this RCT, outcome measures also included quality of life (Q-LES-Q) and cognitive assessments (Shopping List Task [SLT], MMSE, Digital Symbol Substitution Test). Results on these health outcome measures were similar for both drugs. A subgroup analysis of 75 patients 70 years of age or older showed a greater response rate for sertraline than fluoxetine (59 percent vs. 42 percent; $P = 0.027$).[45]

We conducted two meta-analyses of four studies[41,42,44,46] comparing the effects of fluoxetine and sertraline at study endpoint. The outcome of the first meta-analysis was the relative risk (benefit) of being a responder on the HAM-D (improvement of 50 percent or more) at study endpoint (Figure 7).

Pooled results included 940 patients and yielded a statistically significant additional treatment effect for sertraline. The relative risk of being a responder was 1.11 (95% CI, 1.01-1.21) for sertraline relative to fluoxetine. Both random effects and fixed effects models presented similar, statistically significant results. The NNT to gain one additional responder based on the pooled risk difference is 14 (95% CI, 8-22).

The second meta-analysis was an effect size meta-analysis assessing the pooled difference of points on the HAM-D scale (Figure 8). Because of lack of reported data, we limited the analysis to three studies.[41,44,46] We found no statistically significant difference in points on the HAM-D scale between fluoxetine and sertraline. Relative to fluoxetine, sertraline had an additional treatment effect of a 0.75 point reduction in HAM-D (95% CI, -0.45-1.95).

Fluvoxamine vs. paroxetine. One 7-week RCT compared the efficacy and safety of fluvoxamine (50-150 mg/day) and paroxetine (20-50 mg/day) in 60 outpatients with MDD.[50] Results presented no statistically significant differences on HAM-D, HAM-A, CGI, and SCL-56 (Hopkins Symptom Checklist - 56 item). This study did not assess response and remission rates.

Fluvoxamine vs. sertraline. Two 7-week studies compared the depression scores and harms of fluvoxamine (50-150 mg/day) and sertraline (50-200 mg/day).[51,52] One trial was conducted in a mixed (84 percent unipolar, 16 percent bipolar depression) inpatient population.[52]

Figure 7. Relative risk meta-analysis of response rates comparing fluoxetine with sertraline on the HAM-D

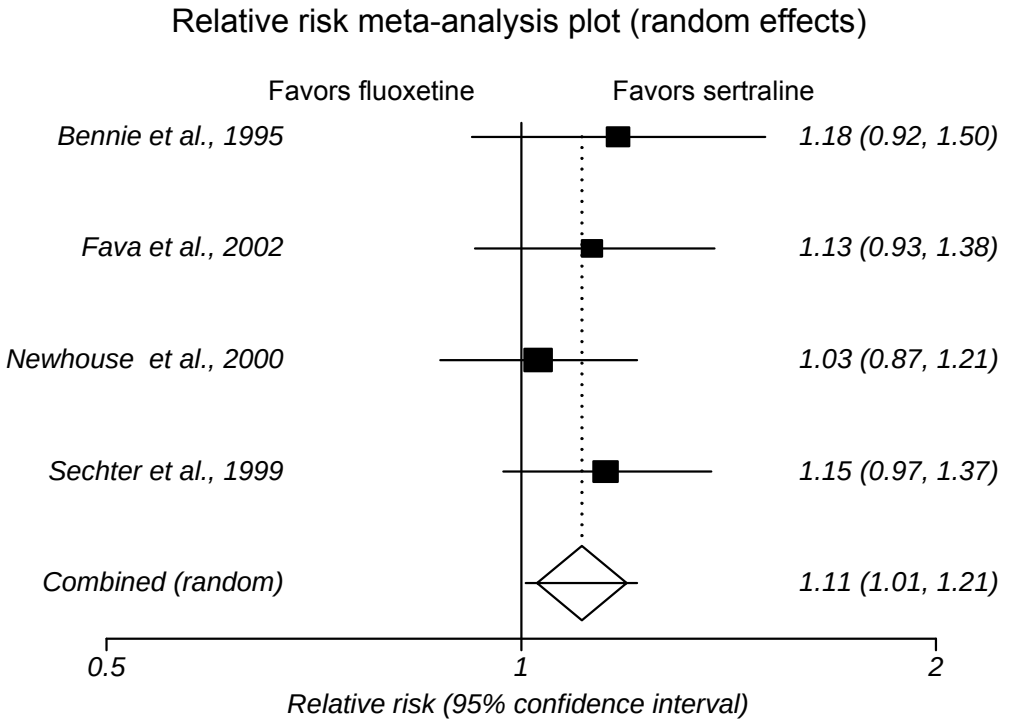
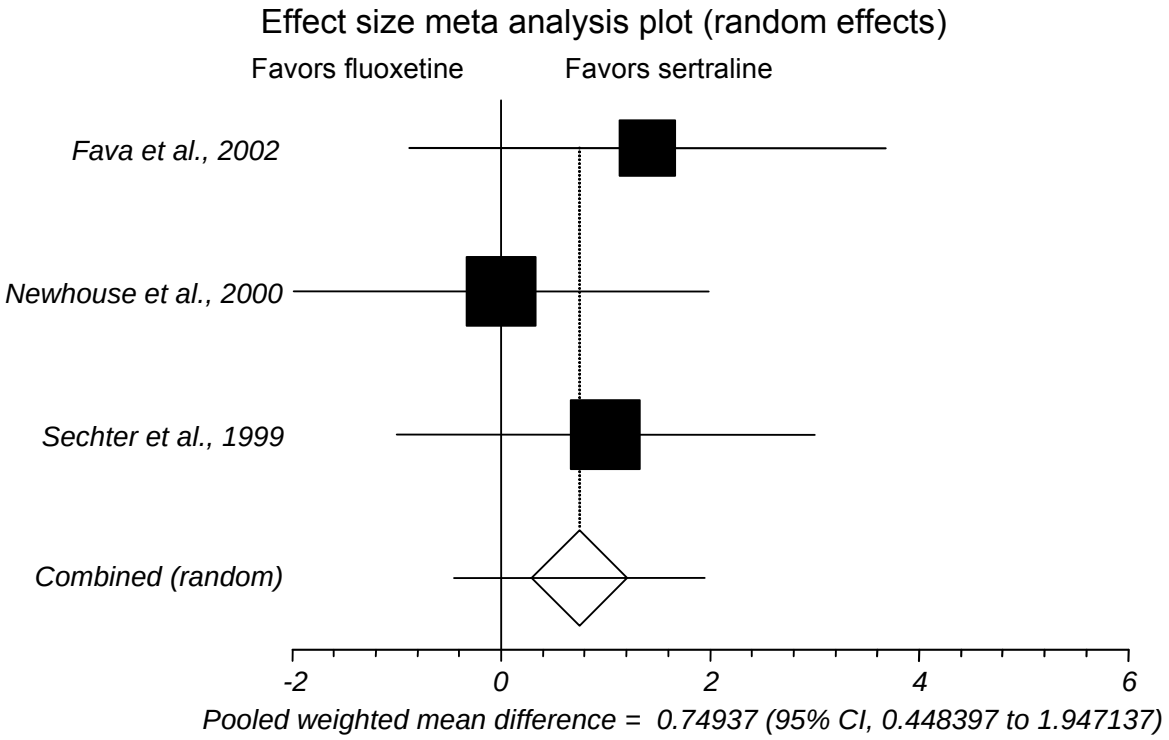


Figure 8. Effect size meta-analysis comparing fluoxetine with sertraline on the HAM-D



In both studies, efficacy did not differ significantly between treatment groups. Both regimens led to significant improvements in depression scores from baseline (HAM-D, CGI). In one study, significantly more patients withdrew because of adverse events in the fluvoxamine group (19 percent) than in the sertraline group (2 percent; $P = 0.016$).[51] Sertraline-treated patients reported a significantly greater rate of sexual dysfunction than patients on fluvoxamine (28 percent vs. 10 percent; $P = 0.047$).

Paroxetine vs. sertraline. A Swedish RCT compared paroxetine (20-40 mg/day) with sertraline (50-150 mg/day) in a 24-week study involving 353 patients.[53] Outcome measures included MADRS, CGI, and Battelle Quality of Life Measure (BQOL). LOCF analysis yielded no significant differences in primary outcome measures (MADRS, CGI) at any point in time. Clinically significant improvement occurred over baseline among all quality of life factors. Treatment groups did not differ significantly on BQOL factors. Diarrhea was more frequent in the sertraline group (35.2 percent vs. 15.2 percent; $P < 0.01$). By contrast, patients in the paroxetine group had higher rates of fatigue (45.8 percent vs. 21.0 percent; $P < 0.01$), decreased libido in females (8.8 percent vs. 1.8 percent; $P < 0.05$), micturition problems (6.2 percent vs. 0.6 percent; $P < 0.05$), and constipation (16.4 percent vs. 5.7 percent; $P < 0.01$).

Head-to-head evidence: SSRIs vs. SSNRIs and SNRIs. ***Citalopram vs. mirtazapine.*** A 8-week European study (n = 270) determined the comparative efficacy of citalopram (20-60 mg/day) and mirtazapine (15-60 mg/day) on depression and anxiety symptoms in a mixed inpatient and outpatient population.[54] At study endpoint, results on efficacy measures (MADRS, HAM-A, CGI-S, Leeds Sleep Evaluation Questionnaire) and a quality of life measure (Q-LES-

Q) were similar between treatment groups. Response rates on MADRS reached 88 percent in the citalopram and 85 percent in the mirtazapine group ($P = 0.54$). Mirtazapine, however, had a faster onset of action with significantly greater response rates on MADRS, HAM-A, CGI-S, and Q-LES-Q at day 14. Mirtazapine led to weight gain in significantly more patients than citalopram (15.3 vs. 4.5 percent; $P < 0.05$); citalopram had a significantly higher rate of nausea than mirtazapine (20 percent vs. 10.2 percent; $P < 0.05$) Overall discontinuation rates because of adverse events did not differ significantly between the two groups.

Citalopram vs. venlafaxine. A 6-month European study compared citalopram (10-30 mg/day) with venlafaxine XR (75-150 mg/day) for the treatment of depression in elderly outpatients (mean age 73 years).[55] No statistical differences in any outcome measures (MADRS, CGI-S, CGI-I) could be detected at study endpoint. The remission rates were 19 percent for venlafaxine and 23 percent for citalopram (P = NR). Both treatment groups reached a 93 percent response rate.

Escitalopram vs. venlafaxine. Two 8-week studies assessed the comparative effectiveness of escitalopram and venlafaxine XR.[56,57] One study assigned 293 patients to escitalopram (10-20 mg/day) or venlafaxine XR (75-150 mg/day).[57] The groups did not differ significantly in response (escitalopram: 77.4 percent; venlafaxine XR: 79.6 percent; P = NR) or remission (escitalopram: 69.9 percent; venlafaxine XR: 69.7 percent; P = NR). Survival analysis of the ITT population indicated that escitalopram-treated patients achieved sustained remission 6.6 days earlier than patients on venlafaxine XR ($P < 0.01$). Significantly fewer patients on escitalopram than on venlafaxine XR reported nausea (17 percent vs. 26 percent; $P < 0.05$), sweating (6 percent vs. 12.5 percent; $P < 0.05$), and constipation (2 percent vs. 6 percent; $P < 0.05$).

The second trial also reported that no statistically significant differences were apparent between escitalopram (20 mg/day) and venlafaxine XR (225 mg/day) in response (61 percent vs. 48 percent; P = NR) and remission rates.[56] Significantly fewer patients in the escitalopram group withdrew because of adverse events (4 percent vs. 16 percent; $P < 0.01$) or reported nausea (24 percent vs. 6 percent; $P < 0.05$). This study, however, compared a medium dose of escitalopram to a high dose of venlafaxine XR. Some differences in adverse events might be attributable to the high, fixed-dose regimen of venlafaxine XR.

Fluoxetine vs. duloxetine. A 8-week RCT assigned 173 patients to duloxetine (40-120 mg/day), fluoxetine (20 mg/day), or placebo.[58] Results revealed no statistically significant differences between fluoxetine and duloxetine in response rates (45 percent vs. 49 percent; $P = 0.39$). Remission rates at study endpoint favored duloxetine but did not reach statistical significance (43 percent vs. 30 percent; $P = 0.82$). However, the fixed-dose design for fluoxetine but not for duloxetine introduces equivalency issues and reduces the validity of this direct comparison.

Fluoxetine vs. mirtazapine. Three trials compared the efficacy of fluoxetine and mirtazapine.[59-61] Two studies enrolled either exclusively[60] or a large percentage[61] of inpatients and outpatients with severe depression (HAM-D > 25). In both of these trials, treatments did not differ on any efficacy measures (MADRS, HAM-D, CGI) or quality of life measures (Q-LES-Q) at endpoint (6 and 8 weeks). Both trials reported a faster onset of mirtazapine but no differences

in remission rates at endpoint. These findings are consistent with results from the third study, which was conducted in Taiwanese outpatients with moderate depression.[59]

In all three studies, patients treated with mirtazapine gained weight; by contrast, those treated with fluoxetine lost weight. In two studies, the differences reached statistical significance.[60,61] In one trial, 10.3 percent of patients in the mirtazapine group experienced an increase in body weight of more than 7 percent from baseline as did 0.9 percent of patients on fluoxetine.[60]

Fluoxetine vs. venlafaxine. Nine studies compared the efficacy of fluoxetine to venlafaxine.[62-70] One study was conducted in inpatient populations.[69] One trial was conducted in outpatients with concomitant anxiety (minimum score of 8 on Covi Anxiety Scale).[64] The studies lasted from 6 weeks to 12 weeks. Except in one study,[70] results consistently presented greater efficacy of venlafaxine than fluoxetine; in three studies, this difference reached statistical significance.[62,64,65]

We conducted a meta-analysis of eight studies comparing fluoxetine to venlafaxine.[62-67,69,70] All studies were financially supported by the manufacturer of venlafaxine. We excluded one study because of missing data.[68] The main outcome measure was the relative risk (benefit) of being a responder on the HAM-D scale at study endpoint.

Results (Figure 9), based on 1,814 patients, present a modest additional treatment effect for venlafaxine, just reaching statistical significance (RR, 1.12; 95% CI, 1.01-1.24). The NNT to achieve one additional responder was 12 (95% CI, 7-50) for the random effects model; the fixed effects model yielded similar significant results. An effect size meta-analysis (Figure 10) yielded a statistically nonsignificant additional reduction of 1.31 points (95% CI, -0.28-2.91) for venlafaxine compared with fluoxetine on the HAM-D scale. The clinical significance of this difference is questionable.

In a sensitivity analysis, we limited studies to those with outpatients only. Results did not differ substantially from findings of analyses that combined inpatient and outpatient subjects. Patients in the venlafaxine group had statistically significantly higher response rates than did patients in the fluoxetine group (RR, 1.12; 95% CI, 1.00-1.25). Again, the additional effect size is modest, just reaching statistical significance.

Figure 9. Relative risk meta-analysis of response rates comparing fluoxetine with venlafaxine on the HAM-D

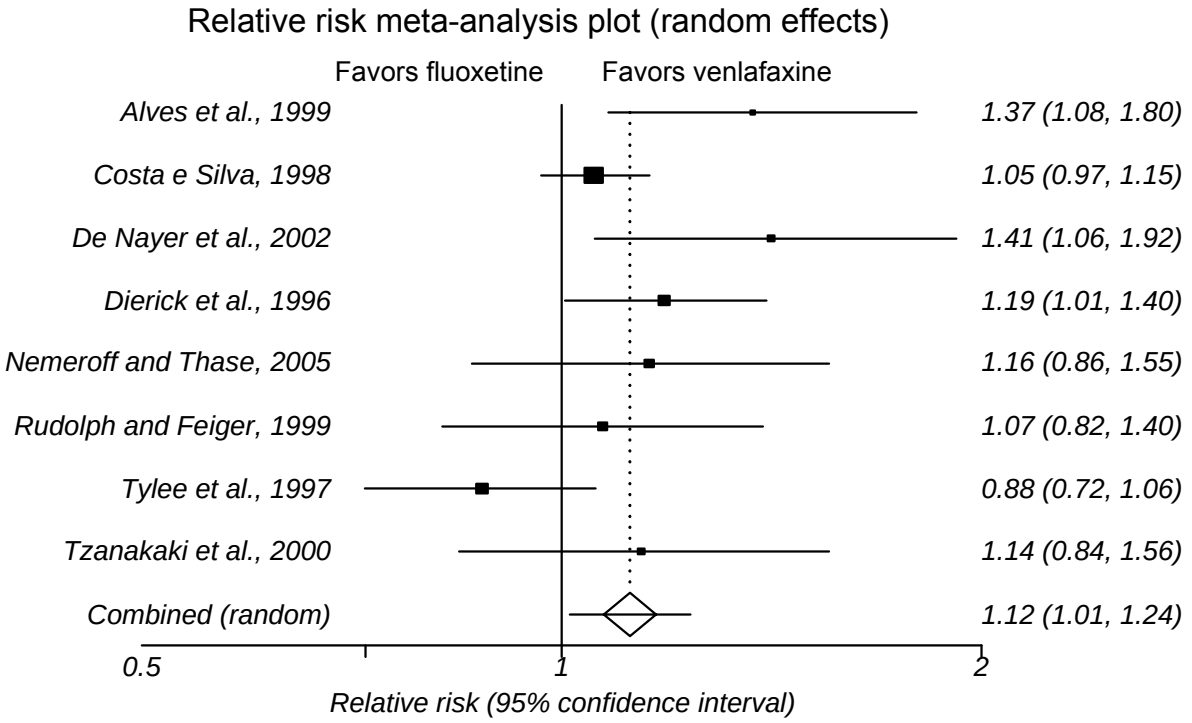
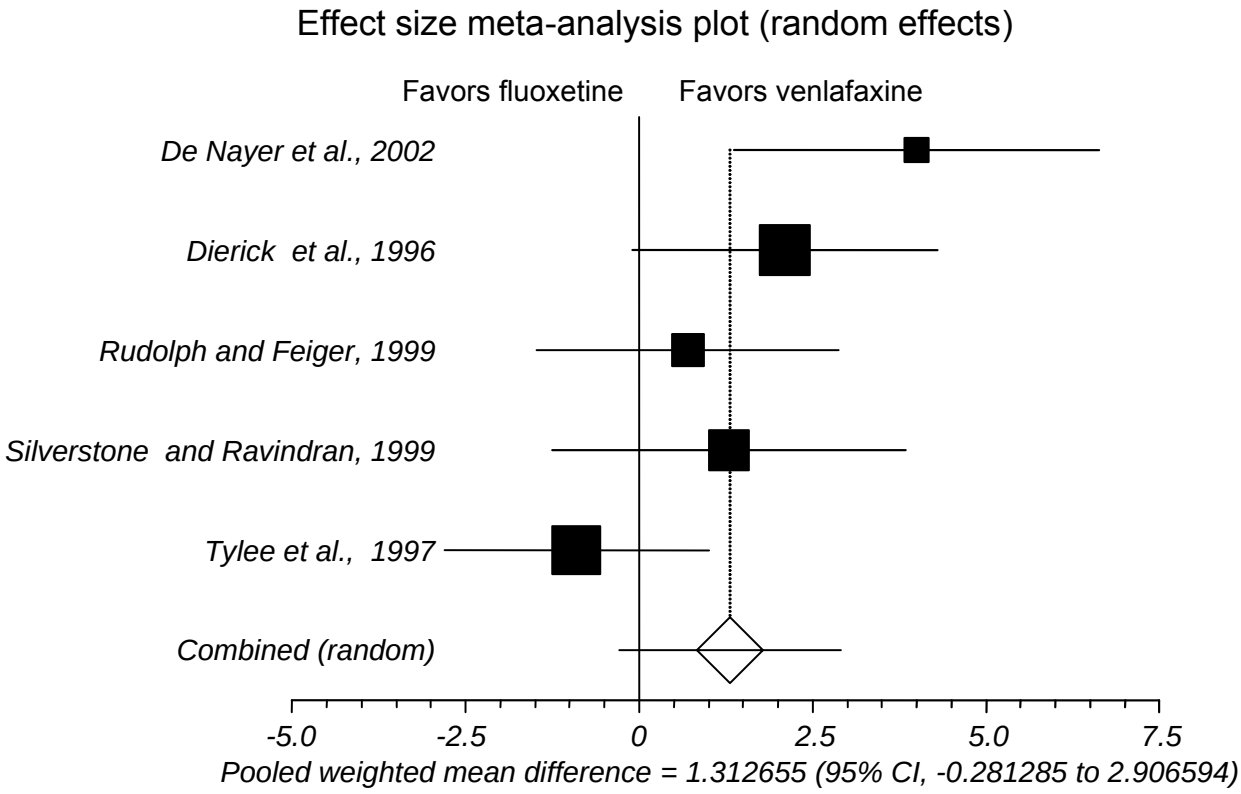


Figure 10. Effect size meta-analysis comparing fluoxetine with venlafaxine on the HAM-D



These findings are consistent with results of a meta-analysis reported by Smith et al.[104] Compared with fluoxetine, venlafaxine yielded a modest but statistically significantly greater standardized effect size (-0.14; 95% CI, -0.22 - -0.06) and a significantly greater odds ratio (OR) for remission (OR, 1.42; 95% CI, 1.17-1.73). The OR for response was numerically greater for venlafaxine but not significant (OR, 1.17; 95% CI, 0.99-1.38).

Paroxetine vs. duloxetine. An 8-week, fixed-dose trial assessed the comparative efficacy of paroxetine (20 mg/day), duloxetine (80 mg/day), duloxetine (120 mg/day), and placebo.[71] These are comparisons between a low-to-medium dose of paroxetine (20 mg) and a medium dose (80 mg) and high dose (120 mg) of duloxetine. Patients in the three active drug groups did not differ significantly in either response (74 percent; 65 percent; 71 percent; P = NR) or remission (44 percent; 46 percent; 52 percent; P = NR). The Patient Global Impression of Improvement (PGI-I) score was significantly better in patients on paroxetine than on 80 mg/day duloxetine.

Paroxetine vs. mirtazapine. Two trials, one conducted in Germany[72] and one in the United States,[73] assessed the efficacy of paroxetine (20-40 mg/day) and mirtazapine (15-45 mg/day). The US study was conducted in depressed patients 65 years or older.[73] In both trials, paroxetine and mirtazapine were equally effective in reducing HAM-D scores at the endpoint. Mirtazapine led to a faster response in both trials. In the German study, 23.2 percent of mirtazapine-treated patients and 8.9 percent of paroxetine-treated patients responded to the treatment at week 1 ($P <$ 0.002). A Kaplan-Meier analysis in one trial showed a significantly faster time to response for mirtazapine than for paroxetine (mean 26 days vs. mean 40 days; $P = 0.016$).[73] No significant difference in response rates on the CGI scale was noted. Both trials reported weight gain in significantly more patients treated with mirtazapine than with paroxetine ($P < 0.05$). Paroxetine-treated patients in the US study reported significantly higher rates of nausea, tremor, and flatulence ($P < 0.05$). The NNT to yield one additional patient responding with mirtazapine at weeks 1 or 2 is 7.

Paroxetine vs. venlafaxine. Two studies compared paroxetine with venlafaxine.[74,75] A Spanish study compared the effects of paroxetine (20-40 mg/day) with venlafaxine (75-150 mg/day) in outpatients (N = 84) with either MDD or dysthymia over 24 weeks.[74] The majority of patients (88 percent) were female. The percentage of dysthymic patients was not reported, and the authors did not differentiate between dysthymia and mild or moderate depression. Loss to followup was 32 percent, with a substantially higher loss to followup in the venlafaxine group (39 percent vs. 26 percent). Response and remission rates favored venlafaxine at all time points. The difference in remission rates reached statistical significance at week 12 (57 percent vs. 33 percent; $P = 0.011$). ITT analysis yielded no significant differences between treatment groups on any primary outcome measures (HAM-D, MADRS, CGI) at 24 weeks.

A British fixed-dose trial lasting 12 weeks randomized 361 mainly moderately ill patients (based on CGI severity score) treated in 43 general practices to either paroxetine (20 mg/day) or venlafaxine XR (75 mg/day).[75] Study groups did not differ significantly in efficacy measures, quality of life scores, or adverse events.

Sertraline vs. mirtazapine. One European study examined the onset of efficacy of sertraline (50-150 mg/day) compared with that of mirtazapine (30-45 mg/day) in 346 outpatients.[76] Onset of action was faster for the mirtazapine group than for the sertraline group on HAM-D and

MADRS. Significantly more patients achieved response and remission on mirtazapine than on sertraline after the first 2 weeks (data not reported in the article; $P < 0.05$) No significant difference could be detected at endpoint. Subgroup analysis in patients with severe depression (HAM-D > 25) led to similar findings. A significantly higher number of patients withdrew because of adverse events in the mirtazapine group (12.5 percent vs. 3 percent; P = NR), and significantly more patients on mirtazapine than on sertraline had an increase in body weight of more than 7 percent (14.6 percent vs. 0 percent; $P = 0.01$).

Sertraline vs. venlafaxine. Two 8-week trials, both rated good quality, compared the efficacy of sertraline to venlafaxine; they yield mixed results regarding differences in efficacy.[77,78] In a Scandinavian study (N = 147), venlafaxine (75-150 mg/day) was significantly more efficacious than sertraline (50-100 mg/day) on the HAM-D (response: 83 percent vs. 68 percent; $P = 0.05$, remission: 68 percent vs. 45 percent; $P = 0.008$).[77] The other study (N = 163) assessed quality of life as the primary outcome measure (Q-LES-Q).[78] LOCF results at 8 weeks did not detect significant differences in quality of life and response and remission rates between treatment groups. Subgroup analyses in this trial, focused on patients with anxious or severe depression, indicated that response and remission rates did not differ significantly between sertraline (50-150 mg/day) and venlafaxine XR (75-225 mg/day).

Head-to-head evidence: SSRIs vs. other second-generation antidepressants. ***Fluoxetine vs. bupropion***. Two trials compared the efficacy and harms of fluoxetine and bupropion.[79,80] Both studies reported similar response rates at endpoint; efficacy measures (changes of HAM-D, HAM-A, CGI-S, CGI-I scores) did not differ significantly. In the larger trial (N = 456), bupropion SR (150-400 mg/day) treatment yielded a higher rate than fluoxetine (20-60 mg/day) of patients achieving remission, but this difference was not significant (47 percent vs. 40 percent; P = NR).[79] From week 1 until endpoint (week 8), significantly more patients on fluoxetine than on bupropion SR were dissatisfied with their overall sexual function (data not reported; $P < 0.05$).

Fluoxetine vs. nefazodone. Two studies with identical protocols examined the effects of antidepressive treatments with either fluoxetine or nefazodone in outpatients with MDD and insomnia.[105,106] Data from these trials and an unpublished study also employing the same protocol were pooled into one analysis.[81]

A total of 125 patients with MDD and sleep disturbance were enrolled for 8 weeks. Effects on sleep were measured by the HAM-D Sleep Disturbance subscale, Inventory for Depressive Symptomatology-Clinician Rated (IDS-C), Inventory for Depressive Symptomatology-Self Rated (IDS-SR), and electroencephalogram measurements. Fluoxetine and nefazodone were similarly efficacious in producing response on the HAM-D scale (45 percent vs. 47 percent; P = NR). Nefazodone led to significantly greater improvements of sleep quality than fluoxetine as assessed by clinician ratings and self-reported evaluations ($P < 0.01$).

Fluoxetine vs. trazodone. Two 6-week trials compared the efficacy and harms of fluoxetine (20-60 mg/day) and trazodone (50-400 mg/day).[82,83] The groups did not differ significantly in any outcome measures (HAM-D, CGI-I, CGI-S, PGI-I). Remission rates in the larger study (N = 126), however, favored fluoxetine over trazodone at study endpoint (51 percent vs. 42 percent;

P = NR).[82] Moreover, significantly fewer patients on fluoxetine than on trazodone experienced sedation or adverse events associated with sedation (22 percent vs. 43 percent; *P* = 0.11)

Paroxetine vs. bupropion. One RCT examined the efficacy of paroxetine (10-40 mg/day) and bupropion SR (100-300 mg/day) in 100 outpatients ages 60 years or older (range 60-88 years) over 6 weeks.[84] Relative to baseline, both groups significantly improved in all outcome measures (HAM-D, HAM-A, CGI-I, CGI-S), but the treatment groups did not differ significantly. Response rates were similar in both groups (paroxetine, 77 percent; bupropion SR, 71 percent; *P* = NR). Both treatment groups improved significantly in quality of life scales (Quality of life in Depression Scale [QLDS], SF-36) between baseline and endpoint (*P* < 0.0001); again, the treatment groups did not differ significantly.[107]

Paroxetine vs. nefazodone. Two studies determined the comparative efficacy of paroxetine and nefazodone on depression and sleep improvement.[85,86] The larger trial enrolled 206 moderately depressed patients to an 8-week, acute-phase trial comparing nefazodone (200-600 mg/day) with paroxetine (20-40 mg/day).[85] Both groups showed significant improvements from baseline HAM-A, HAM-D, and MADRS scores. Response rates were similar for paroxetine and nefazodone (60 percent vs. 58 percent; *P* = NR). The second trial provided consistent results for the comparative antidepressive efficacy.[86] Nefazodone, however, led to significantly greater improvements than paroxetine in objective sleep measures.

Paroxetine vs. trazodone. A European study compared paroxetine (20-40 mg/day) with trazodone (150-400 mg/day) in 108 outpatients with MDD.[87] Study duration was 6 weeks. No differences in any efficacy outcome measures could be detected (HAM-D, CGI-S, CGI-I, MADRS). Response rates (91 percent vs. 87 percent; *P* = NR) and remission rates (68 percent vs. 69 percent; *P* = NR) did not differ significantly between paroxetine and trazodone.

Sertraline vs. bupropion. Three studies compared the efficacy and harms of sertraline and bupropion.[88-90,108] Studies lasted from 8 weeks to 16 weeks. All three studies reported no statistically significant differences in efficacy on any outcome measure (HAM-D, CGI-I, CGI-S, HAM-A). Response rates in the largest trial (N = 364) were 61 percent for sertraline and 66 percent for bupropion SR (*P* = NR).[88]

In all three studies, patients on sertraline had statistically significantly higher rates of sexual dysfunction than patients on bupropion. Two RCTs assessed the incidence of sexual dysfunction during 8 weeks of treatment with sertraline (50-200 mg/day), bupropion SR (150-400 mg/day), or placebo as primary outcome measures using DSM-IV definitions for sexual dysfunction disorders.[88,89] In another study, discontinuation rates because of sexual adverse events were significantly higher in the sertraline group than the bupropion SR group (13.5 percent vs. 3.3 percent, *P* = 0.004).[90] In addition, in this study some adverse events (nausea, diarrhea, somnolence, sweating) were significantly more common among patients treated with sertraline than among those on bupropion SR (*P* < 0.05).

Sertraline vs. nefazodone. A multicenter European study assessed the efficacy and harms of sertraline (50-200 mg/day) and nefazodone (100-600 mg/day) among 160 outpatients with moderate to severe depression.[92] ITT analysis in this 6-week trial did not yield significant differences in efficacy between treatment groups. Response rates were similar between patients

treated with sertraline and those treated with nefazodone (57 percent vs. 59 percent; P = NR). Additional outcome measures assessed by questionnaire were sexual function and satisfaction under antidepressant treatment. Overall satisfaction with sexual function was significantly higher in the nefazodone group ($P < 0.01$). Among men, 67 percent in the sertraline group and 19 percent in the nefazodone group reported difficulty with ejaculation ($P < 0.01$). Other adverse events did not differ significantly between the two groups.

Head-to-head evidence: SNRIs vs. SNRIs. *Mirtazapine vs. venlafaxine.* An 8-week European trial evaluated efficacy and harms in hospitalized, severely depressed patients (mean HAM-D 29.3) with melancholic features.[93] At study endpoint, no significant differences in any efficacy or quality of life measures were apparent (HAM-D, MADRS, CGI-S, Q-LES-Q, QLDS); however, response rates favored mirtazapine over venlafaxine (62 percent vs. 52 percent; P = NR). During the study, significantly fewer patients on mirtazapine than on venlafaxine dropped out because of adverse events (5.1 percent vs. 15.3 percent; $P = 0.037$). Mirtazapine led to weight gain in significantly more patients than did venlafaxine (10.3 percent vs. 5.1 percent; $P < 0.05$). Venlafaxine had significantly lower rates of constipation (17.1 percent vs. 31.1 percent; $P = 0.056$) and sweating (15.8 percent vs. 35.1 percent; $P \leq 0.05$) than venlafaxine.

Head-to-head evidence: SNRIs vs. other second-generation antidepressants. *Mirtazapine vs. trazodone.* Two studies compared mirtazapine with trazodone in patients with MDD.[94,95] One trial was conducted in depressed patients 55 years of age and older;[94] the other was done in hospitalized patients with MDD.[95] Efficacy measures in both trials favored mirtazapine, but differences did not reach statistical significance. In the hospitalized patients, response rates at endpoint were 61 percent for mirtazapine and 51 percent for trazodone (P = NR).[95]

Venlafaxine vs. trazodone. A 6-week study enrolled 225 patients to assess efficacy and harms of venlafaxine (150-400 mg/day), trazodone (75-200 mg/day), and placebo.[96] Efficacy outcomes (HAM-D, MADRS, CGI-S) did not differ significantly between active treatment groups. Response rates at endpoint, however, favored venlafaxine over trazodone (72 percent vs. 60 percent; P = NR). Trazodone led to improvements in sleep disturbance that were statistically significantly superior to those with venlafaxine. Significantly more patients on venlafaxine than on trazodone suffered from nausea (44 percent vs. 19 percent; $P < 0.05$); however, trazodone led to a significantly higher rate of dizziness than venlafaxine (36 percent vs. 17 percent; $P < 0.05$).

Head-to-head evidence: other second-generation antidepressants vs. other second-generation antidepressants. *Bupropion vs. trazodone.* In a two-center study, 124 outpatients were randomly assigned to bupropion (225-450 mg/day) or trazodone (150-450 mg/day).[97] Because of a statistically significant treatment-by-center interaction, the article reported results separately for each center. Overall, in both centers, efficacy results did not differ significantly between the two treatment groups. A postrandomization exclusion rate of 10 percent and an overall loss to followup of 40 percent might compromise the internal validity of this study.

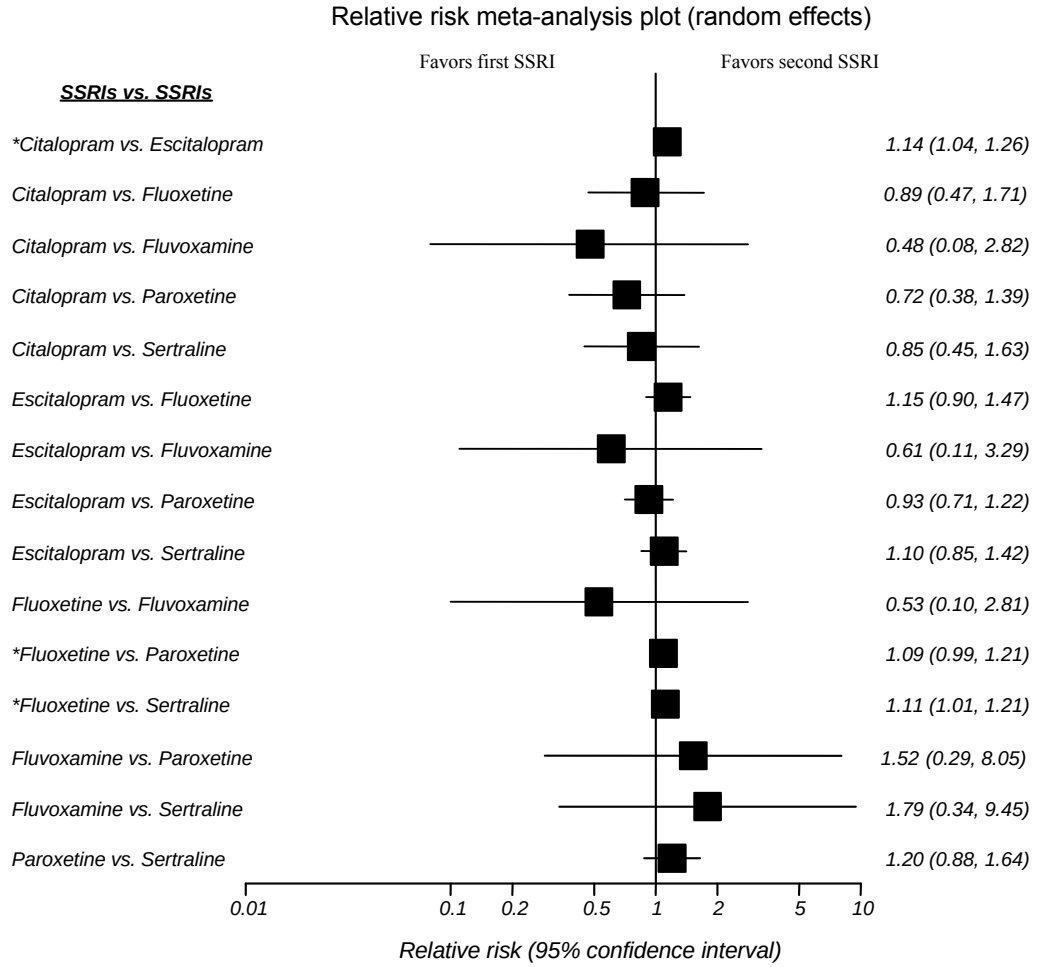
Indirect comparisons. Of 66 possible comparisons, the evidence was sufficient to pool data in meta-analyses for only four comparisons for MDD (those documented in Figures 3 through Figure 10). For the remaining 62 MDD comparisons, we conducted indirect comparisons,

through meta-regression, as outlined in the Methods section. Studies in these meta-regressions can be found in Appendix F; those excluded are listed in Appendix G.

We assessed the relative risk of response to treatment on the HAM-D scale. None of the results of indirect comparisons suggests a statistically significant difference in efficacy between any drugs. However, confidence intervals are often wide and findings do not conclusively demonstrate noninferiority.[109]

In general, findings from indirect comparisons were consistent with available head-to-head studies. Results of direct (denoted by an asterisk) and indirect comparisons are depicted in Figures 11, 12, and 13.

Figure 11. Relative risks of response rates comparing SSRIs with SSRIs on the HAM-D



* Based on meta-analysis of head-to-head trials.

Figure 12. Relative risks of response rates comparing SSRIs, SNRIs, SSNRIs, and other second-generation antidepressants with other second-generation antidepressants on the HAM-D

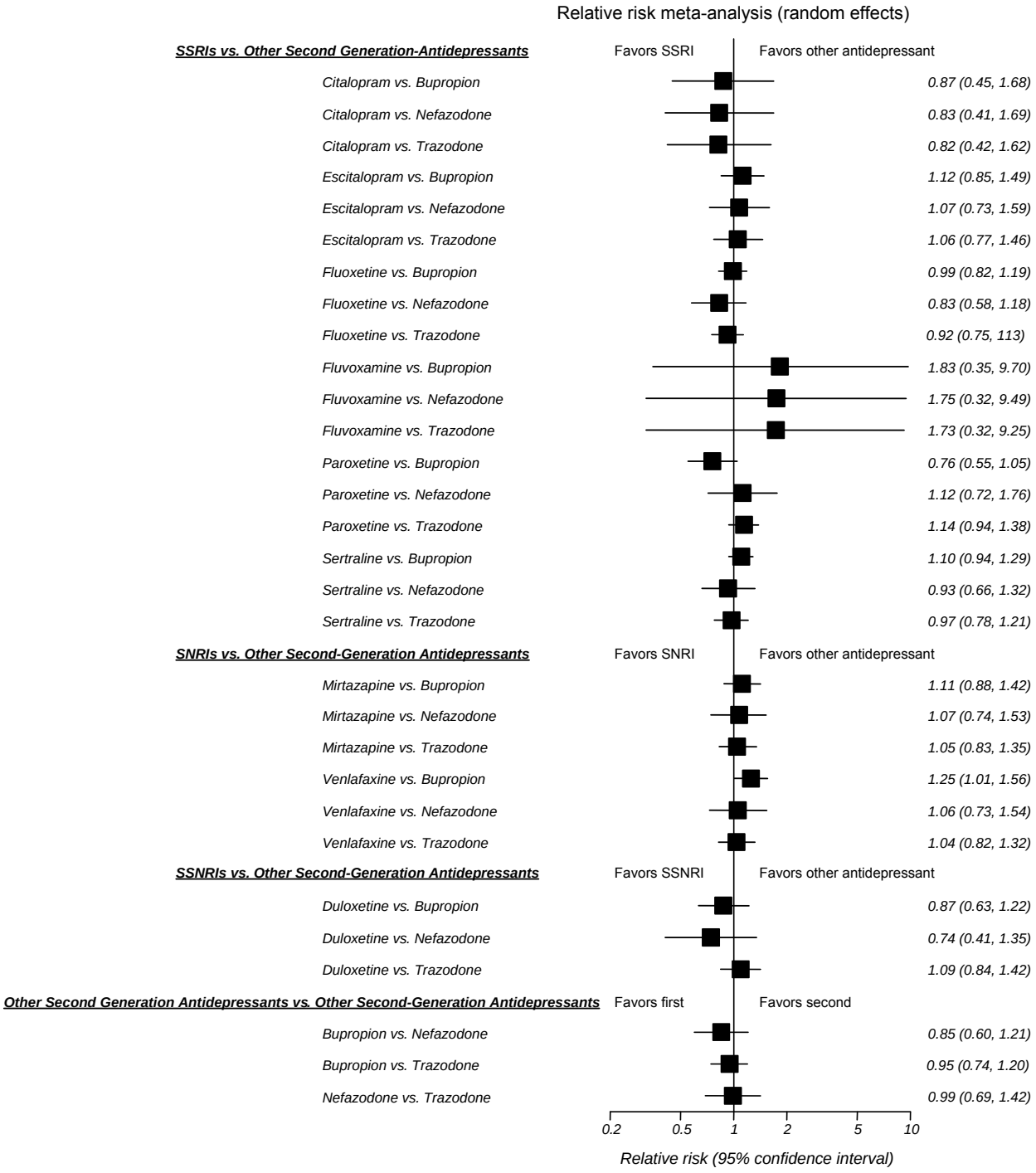
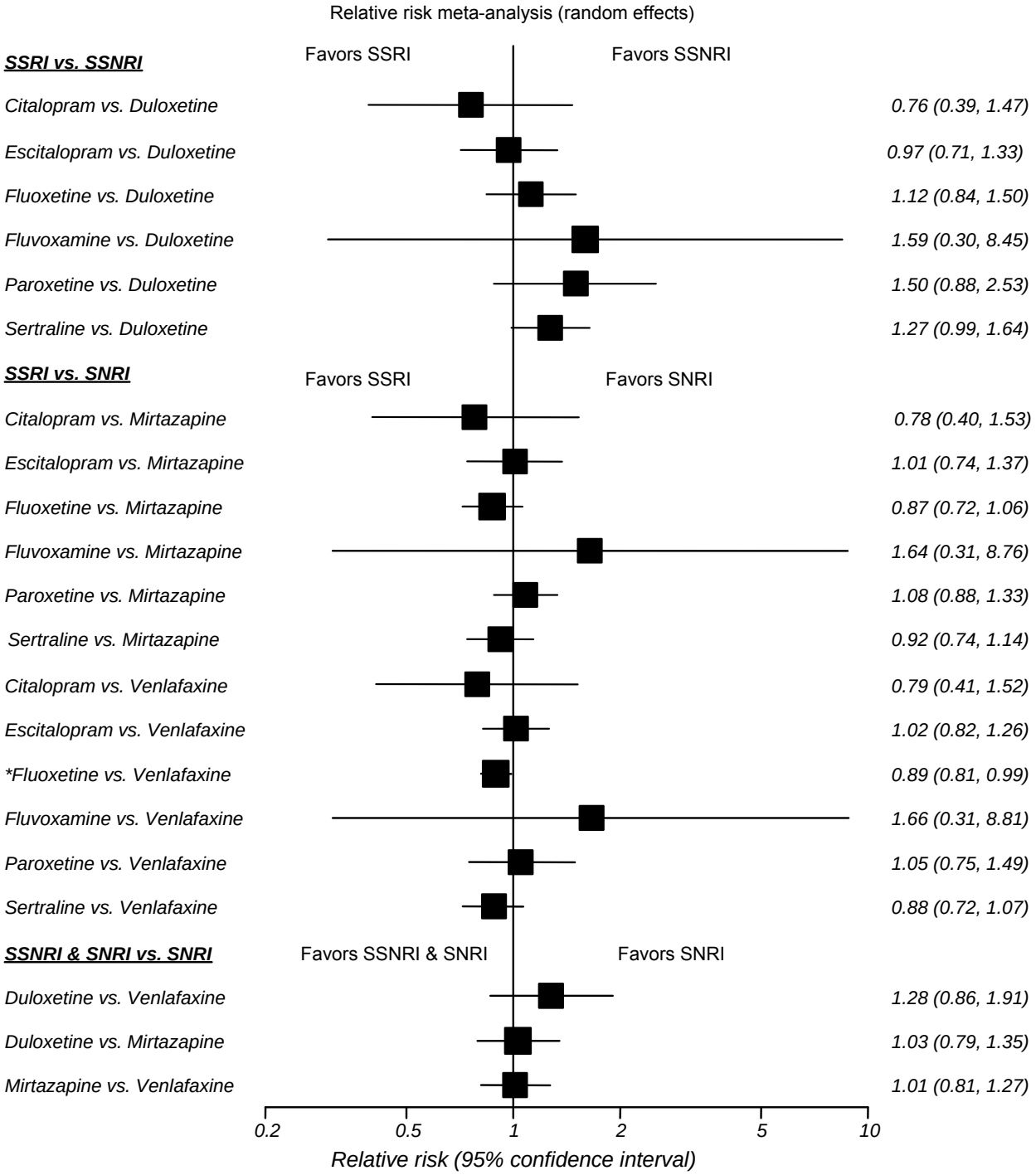


Figure 13. Relative risks of response rates comparing SSRIs with SSNRIs and SSRIs with SNRIs on the HAM-D



*Based on meta-analysis of head-to-head trials.

Dysthymia: Overview

We did not find any head-to-head trials on patients with dysthymia. Five placebo-controlled studies (Table 14) assessed effectiveness, efficacy, and harms of fluoxetine, paroxetine, and sertraline in populations with dysthymia.[99-101,110-113] Four studies were of fair quality; the fifth was of good quality. Details can be found in Evidence Table 1 in Appendix D.

Table 14. Interventions, numbers of patients, results, and quality ratings of studies in adults with dysthymia

Study	Interventions	N	Results	Quality Rating
Devanand et al., 2005[100]	Fluoxetine vs. placebo	90	No difference in response rates and quality of life	Good
Vanelle et al., 1997[101]	Fluoxetine vs. placebo	111	Significantly more responders for fluoxetine	Fair
Barrett et al., 2001[113] Williams et al., 2000[112]	Paroxetine vs. placebo vs. behavioral therapy	656	In patients older than 60 years, significantly greater improvement in symptom scores for paroxetine than for placebo; in patients younger than 60 years, no difference	Fair
Thase et al., 1996[111] Kocsis et al., 1997[110]	Sertraline vs. imipramine vs. placebo	412	Significantly more responders for sertraline than placebo	Fair
Ravindran et al., 2000[99]	Sertraline vs. placebo	310	Significantly more responders and remitters for sertraline	Fair

Dysthymia: Key Points

We identified no head-to-head trials in a population with dysthymia. The significant differences in population characteristics in placebo-controlled trials make the evidence insufficient to identify differences between treatments. The strength of evidence is low.

Five placebo-controlled trials (eight articles) provide mixed evidence on the general efficacy and effectiveness of fluoxetine, paroxetine, and sertraline for the treatment of dysthymia.[99-101,110-113] Specifically:

- Two studies provide mixed evidence about the general efficacy of fluoxetine for the treatment of dysthymia.[100,101]

- One effectiveness study provides mixed evidence on the effectiveness of paroxetine compared with placebo.[112,113] A subgroup of patients older than 60 years showed a significantly greater improvement than those on placebo; a subgroup of patients younger than 60 years did not show any difference in effectiveness between paroxetine and placebo.

- Two studies indicate that sertraline has a significantly greater efficacy in the treatment of dysthymia than placebo.[99,110,111]

Dysthymia: Detailed Analysis

Head-to-head evidence. We identified no head-to head trials.

Placebo-controlled evidence. *Fluoxetine vs. placebo.* Two studies evaluated the efficacy of fluoxetine for treating patients with dysthymia over 12 weeks; the studies provide mixed results.[100,101] An RCT of good quality examined the efficacy and safety of fluoxetine (20-60 mg/day) in patients 60 years of age and older.[100] ITT analysis indicated that fluoxetine had limited efficacy. Response rates on the HAM-D favored fluoxetine over placebo, but the two groups did not differ significantly (27.3 percent vs. 19.6 percent; $P = 0.4$). Likewise, the investigators found no difference in quality of life.

The other study was conducted in patients 18 years of age and older (mean 43 years).[101] Significantly more patients on fluoxetine than on placebo were rated as responders (58.3 percent vs. 35.9 percent; $P = 0.03$). Remission rates favored fluoxetine but did not reach statistical significance (44.4 percent vs. 25.6 percent; $P = 0.07$).

Paroxetine vs. placebo vs. behavioral therapy. A large, primary-care-based effectiveness study randomized 656 patients with dysthymia or minor depression to 11 weeks of paroxetine (10-40 mg/day), placebo, or behavioral therapy.[112,113] Participants were stratified into patients 60 years of age and older (n = 415) and patients younger than 60 years of age (n = 241) for ITT analysis.

In the 60 or older subgroup, paroxetine-treated patients showed a greater change in HSCL-D-20 scores than placebo-treated patients ($P = 0.004$).[112] Effects were similar for patients with dysthymia and minor depression. For older dysthymia patients with high or intermediate baseline functioning scores, paroxetine significantly improved mental health functioning compared with placebo. Overall, however, improvements of mental health functioning were not statistically significantly different between dysthymia patients receiving paroxetine and those receiving placebo.

Among the younger patients, treatment groups did not differ significantly on the HSCL-D-20.[113] For dysthymia only, the remission rate of patients with at least 4 weeks of treatment was significantly higher in the paroxetine group than in the placebo group (80 percent vs. 44 percent; $P = 0.008$). Paroxetine was not more efficacious than placebo in patients with minor depression.

Sertraline vs. placebo. Two RCTs assessed the efficacy of sertraline (50-200 mg/day) for the treatment of dysthymia over 12 weeks.[99,110,111] In both studies, only patients who had had the diagnosis of dysthymia for more than 5 years were eligible; outcomes included quality of life and measures of functional capacity. In both studies, patients on sertraline had significantly greater antidepressant responses than those on placebo (64 percent vs. 44 percent; $P < 0.001$[111] and 51.9 percent vs. 33.8 percent; $P = 0.001$[99]). Likewise, in both studies, sertraline was more efficacious than placebo on psychosocial and quality of life instruments (Global Assessment of Functioning Scale, Social Adjustment Scale [SAD], Quality of Life Enjoyment and Satisfaction Questionnaire [QLSQ], BQOLS).

Subsyndromal Depressive Disorders: Overview

We found no head-to-head RCTs on patients with subsyndromal depressive disorders. The only head-to-head evidence was a nonrandomized, single-blinded trial comparing citalopram with sertraline.[114] Because of the lack of head-to-head evidence, we briefly summarize this study (Table 15), although it did not meet eligibility criteria. In addition, two placebo-controlled studies, both rated fair quality, assessed the efficacy and tolerability of fluoxetine[115] and paroxetine[112,113] in patients with dysthymia (Table 15). Details can be found in Evidence Table 1 in Appendix D.

Table 15. Interventions, numbers of patients, results, and quality ratings of studies in adults with subsyndromal depressive disorders

Study	Interventions	N	Results	Quality Rating
Rocca et al., 2005[114]	Citalopram vs. sertraline	138	No difference	NA
Judd et al., 2004[115]	Fluoxetine vs. placebo	162	Greater improvements on depression scales for fluoxetine than for placebo; no difference in psychosocial outcomes	Fair
Barrett et al., 2001[113] Williams et al., 2000[112]	Paroxetine vs. placebo vs. behavioral therapy	656	In patients older than 60 years, significantly greater improvement in symptom scores for paroxetine than for placebo; in patients younger than 60 years, no difference	Fair

NA, not applicable.

Subsyndromal Depressive Disorders: Key Points

We identified no head-to-head RCTs in a population with subsyndromal depression. A nonrandomized, open-label trial did not detect any differences in efficacy between citalopram and sertraline.[114]

In placebo-controlled trials, significant differences in population characteristics make the evidence insufficient to identify differences between treatments.[112,113,115] In one effectiveness study, effectiveness did not differ significantly between paroxetine and placebo for the treatment of minor depression.[112,113] The strength of evidence is low.

Subsyndromal Depressive Disorders: Detailed Analysis

Head-to-head evidence. We did not find any head-to-head RCTs. A nonrandomized, single-blinded trial (n = 138) lasting 1 year assessed the comparative efficacy and safety of citalopram and sertraline in patients with late-life minor depression or other subsyndromal depressive disorders.[114] Overall, both treatments improved depressive symptoms. No significant differences in efficacy could be detected at any time point. At the end of the study, remission was achieved by 53 percent of patients on citalopram and 42 percent on sertraline ($P = 0.25$). Likewise, no differences in psychosocial functioning emerged.

Placebo-controlled evidence. Two studies were conducted in populations with minor depression.

Fluoxetine vs. placebo. A 12-week trial (N = 162) evaluated the efficacy of fluoxetine in patients with minor depression.[115] Improvements on depression scales (HAM-D, Beck Depression Inventory [BDI], IDS-C) were statistically significantly greater for patients receiving fluoxetine than for those receiving placebo. Likewise, the overall severity of illness (CGI-S) improved statistically significantly more in the fluoxetine than in the placebo group ($P = 0.002$). No significant differences could be detected in psychosocial outcomes.

Paroxetine vs. placebo. A large primary-care-based effectiveness study randomized 656 patients with dysthymia or minor depression to 11 weeks of paroxetine (10-40 mg/day), placebo, or behavioral therapy.[112,113] Participants were stratified into patients 60 years and older (n = 415) and patients younger than 60 years (n = 241) for ITT analysis.

In the 60 or older subgroup, patients receiving paroxetine showed a greater change in HSCL-D-20 scores than those receiving placebo ($P = 0.004$), but those on paroxetine did not demonstrate more change than patients on behavioral therapy ($P = 0.17$).[112] Effects were similar for patients with dysthymia and minor depression. Paroxetine was not more efficacious than placebo in patients with minor depression in the younger subgroup.[113]

Key Question 2: Efficacy or effectiveness for maintaining remission or for treating patients with unresponsive or recurrent disease

This section deals with two issues on efficacy or effectiveness of medications:

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

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

Overview

In all, we had 27 studies relating to these two questions (Table 16). Three head-to-head RCTs compared the efficacy of one second-generation antidepressant with another for preventing relapse or recurrence.[47,96,116,117] Comparisons included fluoxetine vs. sertraline,[47] fluvoxamine vs. sertraline,[116,117] and trazodone vs. venlafaxine (shown in italics in Table 16).[96] Another 21 RCTs[118-138] provide additional placebo-controlled evidence to support the general efficacy of bupropion, citalopram, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone,

paroxetine, sertraline, trazodone, and venlafaxine for maintaining remission in patients with depressive disorders (Table 16). No trial assessed the efficacy of duloxetine for preventing relapse or recurrence. Two effectiveness studies[139,140] and one efficacy trial[141] compared one second-generation antidepressant with another for treating patients who had not responded or could not tolerate at least one previous treatment.

Using the management framework depicted in the introduction (Figure 1), we characterized studies that assessed continuation treatment of patients who had responded or remitted with acute-phase treatment as *relapse prevention* studies. Relapse prevention studies typically included an open-label, acute-phase treatment and a double-blind, randomized, placebo-controlled continuation-phase treatment. The duration of continuation treatment in these trials ranged from 14 weeks to 72 weeks.

We further denoted studies that assessed maintenance treatment among patients who had remained in remission following acute and continuation treatment as *recurrence prevention* studies. These studies usually included an open-label acute phase, then an open-label continuation phase for acute-phase responders, followed by a randomized, double-blind, placebo-controlled maintenance phase for patients who had not relapsed. The maintenance phase in these trials lasted from 36 weeks to 100 weeks.

Table 16. Number of head-to-head comparisons and placebo-controlled studies for assessment of relapse and recurrence

Relapse Prevention (Continuation Treatment ≤ 9 months)		Recurrence Prevention (Maintenance Treatment)	
Comparison	**Number of Studies**	**Comparison**	**Number of Studies**
Head-to-Head Trials			
Fluoxetine vs. sertraline	1	*Fluvoxamine vs. sertraline*	1
		Trazodone vs. venlafaxine	1
Placebo-Controlled Trials			
Bupropion vs. placebo	1	Bupropion vs. placebo	0
Citalopram vs. placebo	2	*Citalopram vs. placebo*	2
Duloxetine vs. placebo	0	Duloxetine vs. placebo	0
Escitalopram vs. placebo	1	Escitalopram vs. placebo	0
Fluoxetine vs. placebo	2	*Fluoxetine vs. placebo*	1
Fluvoxamine vs. placebo	0	*Fluvoxamine vs. placebo*	1
Mirtazapine vs. placebo	1	Mirtazapine vs. placebo	0
Nefazodone vs. placebo	1	*Nefazodone vs. placebo*	1
Paroxetine vs. placebo	1*	*Paroxetine vs. placebo*	2*
Sertraline vs. placebo	1	*Sertraline vs. placebo*	3
Trazodone vs. placebo	0	*Trazodone vs. placebo*	1†
Venlafaxine vs. placebo	1	*Venlafaxine vs. placebo*	2†

* One trial reported continuation-phase and maintenance-phase results.

† Includes placebo comparison from a head-to-head trial of trazodone and venlafaxine.

Studies that compared one second-generation antidepressant with another in treating resistant depressive disorders varied in design, although all studies randomized patients with MDD, dysthymia, or minor depression to an alternative treatment after they had either failed or could

not tolerate a previous treatment. We characterized patients with treatment-resistant depressive disorder as those who had failed initial acute-phase treatment. We searched for studies specifically assessing treatment of relapse (i.e., loss of response during continuation treatment) and recurrence (i.e., a new distinct episode), but studies addressing this population included relapsing patients and/or those with recurrent depression among those studied in acute-phase treatment trials (i.e., KQ 1).

Investigators generally determined the initial inclusion of patients on a criteria-based diagnosis (e.g., DSM-III-R, DSM-IV) and a predefined cutoff point of a universally used depression scale (e.g., HAM-D ≥ 18 or MADRS ≥ 19). Subsequent inclusion criteria varied. Some studies randomized patients who had demonstrated a clinically significant response to open-label treatment (e.g., ≥ 50 percent improvement from baseline on the HAM-D or MADRS). Other studies used a predefined cutoff point on a depression scale to identify and randomize those who were in remission (e.g., HAM-D ≤ 9, MADRS ≤ 12, CGI-I ≤ 2). Most studies assessed relapse or recurrence using a predefined cutoff point on a depression rating scale (e.g., HAM-D > 18, MADRS > 19, CGI-S ≥ 4), but the specific cutoff point varied widely.

Because most studies received a fair rating for quality (internal validity), we denote quality in this section only for those rated good or poor. Trial reporting was often incomplete. Most articles did not report their methods of randomization or allocation concealment. Even though investigators frequently used ITT analysis, few authors reported the overall number of patients lost to followup from randomization to the end of the trial. Because some studies defined reasons for attrition such as withdrawals because of lack of efficacy or adverse events as endpoints, we did not use loss to followup as an exclusion criteria for these studies.

Data were insufficient to use placebo-controlled trials for making indirect comparisons between drugs.

Detailed information on all studies reviewed for KQs 2a and 2b can be found in Evidence Table 2 in Appendix D.

Maintaining Remission: Key Points

In three head-to-head trials,[47,96,116,117] the overall efficacy for maintaining remission does not differ between fluoxetine and sertraline,[47] fluvoxamine and sertraline,[116,117] and trazodone and venlafaxine.[96] We rated the strength of head-to-head evidence as moderate.

Ten placebo-controlled relapse prevention trials provide consistent efficacy evidence in favor of active treatment over placebo.[118-127,142] Eleven placebo-controlled recurrence prevention trials provide consistent evidence in favor of active treatment over placebo.[119,128-138] Effect sizes generally were consistent across drugs in placebo-controlled efficacy trials. One placebo-controlled recurrence prevention study[136] rated good quality provides general evidence of the effectiveness of sertraline for maintaining remission in patients with recurrent major depression.

Maintaining Remission: Detailed Analysis

Head-to-head evidence. Three head-to-head trials compared one second-generation antidepressant with another for maintaining remission (Table 17).[47,96,116,117]

Table 17. Head-to-head studies of relapse prevention and recurrence prevention

Study	Phase	Duration (Weeks)	N	Comparison and Dose (mg/d)	Relapse or Recurrence N (%)		Quality Rating
Van Moffaert et al., 1995[47]	Acute	8	82	Fluoxetine 20-40	NA		Fair
			83	Sertraline 50-100	NA		
	Continuation	24	56	Fluoxetine 20-40	7 (13)	*P* = NR (ns)	
			49	Sertraline 50-100	5 (10)		
Franchini et al., 1997[116]	Acute	NR	NR	NR	NA		Fair
	Continuation	16	NR	NR	NA		
Franchini et al., 2000[117]	Maintenance (2 years)[116]	104	32	Fluvoxamine 200	6 (19)	*P* = 0.88	
			32	Sertraline 100	7 (22)		
	Maintenance (4 years)[117]	208	25	Fluvoxamine 200	5 (20)	*P* = 0.92	
			22	Sertraline 100	3 (14)		
Cunningham et al., 1994[96]	Acute	6	77	Trazodone 150-400	NA		Fair
			72	Venlafaxine 75-200	NA		
			76	Placebo	NA		
	Continuation/ Maintenance	52	30	Trazodone 150-400	4 (13)	*P* = NR (ns)	
			37	Venlafaxine 75-200	3 (8)		
			29	Placebo	4 (14)		

NA, not applicable; NR, not reported; ns, not statistically significant.

Fluoxetine vs. sertraline. One trial compared the efficacy of fluoxetine and sertraline for preventing relapse during a 24-week continuation phase.[47] A total of 165 patients with major depression were randomized to fluoxetine 20-40 mg/day or sertraline 50-100 mg/day. At 8 weeks, 56 responders (≥ 50 percent reduction in HAM-D or MADRS) in the fluoxetine group and 49 responders in the sertraline group entered the continuation phase, continuing the same dose attained at the end of the acute phase. Relapse rates were similar in the two groups (13 percent and 10 percent, respectively; *P* = NR), but this design may be prone to bias and confounding because patients had not been rerandomized at the start of the continuation phase.

Fluvoxamine vs. sertraline. One trial compared the efficacy of fluvoxamine and sertraline for maintaining remission over 2 years[116] and 4 years.[117] This relatively small Italian study included 64 patients with recurrent depression. After at least 4 months of remission with tricyclic antidepressants (N = 49), SSRIs (N = 4), monoamineoxidase inhibitors (N = 2), or combination treatment (N = 9), investigators randomized patients to fluvoxamine 200 mg/day or sertraline 100 mg/day and followed them for up to 4 years. Recurrence rates (HAM-D > 15) for fluvoxamine and sertraline were similar at 2 years (19 percent vs. 22 percent, respectively; *P* = 0.88) and 4 years (20 percent vs. 14 percent, respectively; *P* = 0.92).

Trazodone vs. venlafaxine. One randomized trial compared the efficacy and safety of trazodone and venlafaxine over a 1-year continuation/maintenance phase.[96] A total of 225 patients with major depression received acute treatment with trazodone 150-400 mg/day (N = 77), venlafaxine 75-200 mg/day (N = 72), or placebo (N = 76). After 6 weeks, 30 responders in the trazodone group and 37 in the venlafaxine group (CGI-I score of 1 or 2) were allowed to continue into the long-term phase. Relapse rates were similar in the three groups (13 percent, 8 percent, and 14 percent, respectively; *P* = NR). Fewer patients treated with venlafaxine than with

either trazodone ($P \leq 0.05$) or placebo (P = NR) withdrew from treatment for any reason; this difference reached statistical significance during the long-term phase.

Placebo-controlled evidence. Ten placebo-controlled trials (11 publications) assessed relapse prevention[118-127,142] and 11 trials (12 publications) assessed recurrence prevention.[119,128-138] Because the duration of acute, continuation, and maintenance phase treatment is not consistent in all patients, and because the definition of these treatment phases is not universal, some studies described below (Table 18) can be categorized as addressing both relapse and recurrence prevention.

Bupropion vs. placebo. One trial assessed relapse prevention with bupropion.[118] Patients with recurrent major depression (N = 816) were treated openly for 8 weeks with bupropion SR 300 mg/day. Those who responded (CGI-I score of 1 or 2 during the last 3 weeks of the acute phase) were randomized to placebo (N = 213) or continuation treatment with the same dose of bupropion SR (N = 210). After 44 weeks, relapse rates were statistically significantly lower for patients on bupropion than for those on placebo (37 percent vs. 52 percent, respectively; P = 0.004). The median time to relapse, as defined by the need for treatment intervention after randomization into the double-blind phase, was 24 weeks for placebo and at least 44 weeks for bupropion.

Citalopram vs. placebo. Two trials assessed relapse prevention[120,143] and two other trials assessed recurrence prevention.[128,129] Both relapse prevention trials randomized patients who responded in the acute phase (MADRS ≤ 12) to placebo or continuation treatment with citalopram. Statistically significantly fewer patients on citalopram than on placebo relapsed after 24 weeks in both trials. Relapse rates were 14 percent and 24 percent, respectively (P = 0.04), in one trial, and 11 percent (pooled) and 31 percent, respectively ($P < 0.02$), in the other trial. Both recurrence prevention trials included open-label, acute-phase treatment with citalopram 20-60 mg/day (6 weeks to 9 weeks), followed by 16 weeks of open-label continuation treatment at the same dose for responders (MADRS ≤ 11).[128,129] Patients who had not relapsed (MADRS ≤ 22) during the continuation phase were randomized to 48 weeks of double-blind maintenance treatment with citalopram or placebo. Recurrence rates were lower for citalopram-treated patients than for placebo-treated patients in both trials (18 percent vs. 43 percent, respectively; $P < 0.001$,[128] and 32 percent vs. 67 percent, respectively; P = NR[129]).

Escitalopram vs. placebo. One trial treated MDD patients (N = 502) openly with escitalopram 10-20 mg/day for 8 weeks.[121] Patients who responded (MADRS ≤ 12) were randomized to 36 weeks of double-blind continuation treatment with escitalopram (N = 181) or placebo (N = 93). Relapse rates (MADRS ≥ 22) were statistically significantly lower for escitalopram-treated patients than for placebo-treated patients (26 percent vs. 40 percent, respectively; P = 0.01), and the time to depressive relapse was significantly longer in patients who received escitalopram than in patients who received placebo (P = 0.013).

Fluoxetine vs. placebo. Two trials (three publications) assessed relapse prevention,[122,123,142] and one trial assessed recurrence prevention.[130] Of the relapse prevention studies, one trial sought to determine the optimal length of continuation treatment by randomizing patients who were in remission (HAM-D < 7 for 3 consecutive weeks) during 12 weeks to 14 weeks of acute-

phase treatment with fluoxetine 20 mg/day to 14 weeks, 38 weeks, or 50 weeks of continuation treatment with fluoxetine or placebo.[122] Relapse rates were significantly lower for fluoxetine-treated patients than for placebo-treated patients at 14 weeks (26 percent vs. 49 percent, respectively; $P < 0.001$) and 38 weeks (9 percent vs. 23 percent, respectively; $P = 0.04$), but not at 50 weeks (11 percent vs. 16 percent, respectively; $P = 0.54$). The second trial openly treated 932 patients with MDD for 13 weeks with fluoxetine.[123,142] Responders (HAM-D ≤ 9 and CGI-I ≤ 2) were randomized to 25 weeks of continuation treatment with fluoxetine 20 mg/day (N = 189), fluoxetine 90 mg/week (N = 190), or placebo (N = 122). Relapse rates were statistically significantly lower for both the daily and the weekly doses of fluoxetine than for placebo (26 percent and 37 percent vs. 50 percent, respectively; $P < 0.01$ for placebo comparisons).

The recurrence prevention trial randomized patients who continued to meet remission criteria (HAM-D ≤ 8) during a 6-month continuation period to 1 year of double-blind maintenance treatment with either fluoxetine 20 mg/day (N = 70) or placebo (N = 70).[130] Recurrence rates were statistically significantly lower for fluoxetine-treated patients than for placebo-treated patients (20 percent vs. 40 percent, respectively; $P = 0.01$).

Fluvoxamine vs. placebo. One trial assessed recurrence prevention with fluvoxamine 100 mg/day.[131] Of 436 patients with major depression treated openly with fluvoxamine for 6 weeks, 283 responders (MADRS < 10 and CGI-I ≤ 2) entered 18 weeks of continuation treatment. Patients who sustained their response (MADRS < 12 and no CGI-I score > 2) were randomized to 1 year of double-blind treatment with fluvoxamine (N = 110) or placebo (N = 94). Recurrence rates were statistically significantly lower for fluvoxamine-treated patients than for placebo-treated patients (13 percent vs. 35 percent, respectively; $P < 0.001$).

Mirtazapine vs. placebo. One trial of relapse prevention openly treated patients with recurrent or chronic major depression (N = 410) with mirtazapine 15-45 mg/day for 8 weeks to 12 weeks.[124] Those in remission (HAM-D ≤ 7 and CGI-I ≤ 2) were randomized to 40 weeks of continuation treatment with mirtazapine (N = 76) or placebo (N = 80). Relapse rates were statistically significantly lower for mirtazapine-treated patients than for placebo-treated patients (20 percent vs. 44 percent, respectively; $P = 0.001$).

Table 18. Placebo-controlled studies of relapse prevention and recurrence prevention

Study	Phase	Duration (Weeks)	N	Comparison and Dose (mg/d)	Relapse or Recurrence N (%)		Quality Rating
Weihs et al., 2002[118]	Acute	8	816	Bupropion SR 300	NA		Fair
	Continuation	44	210	Bupropion SR 300	78 (37)	$P = 0.004$	
			213	Placebo	111 (52)		
Hochstrasser et al., 2001[128]	Acute	6-9	427	Citalopram 20-60	NA		Fair
	Continuation	16	327	Citalopram 20-60	NA		
	Maintenance	48	132	Citalopram 20-60	24 (18)	$P < 0.001$	
			137	Placebo	59 (43)		
Klysner et al., 2002[129]	Acute	8	230	Citalopram 20-40	NA		Fair
	Continuation	16	172	Citalopram 20-40	NA		
	Maintenance	48	60	Citalopram 20-40	19 (32)	P = NR	
			61	Placebo	41 (67)		

Montgomery et al., 1992[144]	Acute	6	NR	Citalopram 20-40	NA		Fair
	Continuation	24	48	Citalopram 20	4 (8)	$P < 0.02$*	
			57	Citalopram 40	7 (12)		
			42	Placebo	13 (31)		
Robert and Montgomery, 1995[120]	Acute	8	391	Citalopram 20-60	NA		Fair
	Continuation	24	152	Citalopram 20-60	21 (14)	$P = 0.04$	
			74	Placebo	18 (24)		
Rapaport et al., 2004[121]	Acute	8	502	Escitalopram 10-20	NA		Fair
	Continuation	36	181	Escitalopram 10-20	47 (26)	$P = 0.01$	
			93	Placebo	37 (40)		
Gilaberte et al., 2001[130]	Acute	8	253	Fluoxetine 20-40	NA		Fair
	Continuation	24	179	Fluoxetine 20-40	NA		
	Maintenance	52	70	Fluoxetine 20	14 (20)	$P = 0.01$	
			70	Placebo	28 (40)		
Schmidt et al., 2000[123] Dinan, 2001[142]	Acute	13	932	Fluoxetine 20	NA		Fair
	Continuation	25	189	Fluoxetine 20	49 (26)	$P < 0.01$*	
			190	Fluoxetine 90/week	70 (37)		
			122	Placebo	61 (50)		
Reimherr et al., 1998[122] Michelson et al., 1999[145]	Acute	12-14	839	Fluoxetine 20	NA		Fair
	Continuation	14	299	Fluoxetine 20	77 (26)	$P < 0.001$	
			95	Placebo	46 (49)		
	Continuation	38	105	Fluoxetine 20	9 (9)	$P < 0.04$	
			52	Placebo	12 (23)		
	Continuation	50	28	Fluoxetine 20	3 (11)	$P = 0.54$	
			34	Placebo	6 (16)		
Terra and Montgomery, 1998[131]	Acute	6	436	Fluvoxamine 100	NA		Fair
	Continuation	18	283	Fluvoxamine 100	NA		
	Maintenance	52	110	Fluvoxamine 100	14 (13)	$P < 0.001$	
			94	Placebo	33 (35)		
Thase et al., 2001[124]	Acute	8-12	410	Mirtazapine 15-45	NA		Fair
	Continuation	40	76	Mirtazapine 15-45	15 (20)	$P = 0.001$	
			80	Placebo	35 (44)		
Gelenberg et al., 2003[132]	Acute	12	681	Nefazodone 300-600	NA		Fair
	Continuation	16	269	Nefazodone 300-600	NA		
	Maintenance	52	76	Nefazodone 300-600	23 (30)	$P = 0.043$	
			84	Placebo	40 (48)		
Feiger et al., 1999[125]	Acute	16	467	Nefazodone 400-600	NA		Fair
	Continuation	36	65	Nefazodone 400-600	1 (2)	$P = 0.009$	
			66	Placebo	12 (18)		
Claghorn and Feighner, 1993[133]	Acute	6	240	Paroxetine 10-50	NA		Fair
			237	Imipramine 65-275	NA		
			240	Placebo	NA		
	Continuation	52	94	Paroxetine 10-50	11 (12)	P = NR	
			79	Imipramine 65-275	3 (4)		
			46	Placebo	10 (22)		

Montgomery and Dunbar, 1993[119]	Acute	8	172	Paroxetine 20-30	NA		Fair
	Continuation	16	68	Paroxetine 20-30	2 (3)	*P* < 0.01	
			67	Placebo	13 (19)		
	Maintenance	36	66	Paroxetine 20-30	9 (14)	*P* < 0.05	
			54	Placebo	16 (30)		
Lepine et al., 2004[136]	Remission Stability	8	371	Placebo	NA		Good
	Maintenance	72	189	Sertraline 50-100	32 (17)	*P* = 0.002	
			99	Placebo	33 (33)		
Doogan and Caillard, 1992[126]	Acute	8	480	Sertraline 50-200	NA		Fair
	Continuation	44	185	Sertraline 50-200	24 (13)	*P* < 0.001	
			110	Placebo	48 (46)		
Keller et al., 1998[134] Kocsis et al., 2002[135]	Acute	12	426	Sertraline 50-200	NA		Fair
	Continuation	16	209	Sertraline 50-200	NA		
	Maintenance	76	77	Sertraline 50-200	5 (6)	*P* = 0.002	
			84	Placebo	19 (23)		
Wilson et al., 2003[137]	Acute	8	318	Sertraline 50-200	NA		Fair
	Continuation	16-20	254	Sertraline 50-200	NA		
	Maintenance	100	56	Sertraline 50-100	25 (45)	*P* = 0.21	
			57	Placebo	31 (54)		
Montgomery et al., 2004[138]	Acute/ Continuation	26	495	Venlafaxine 100-200	NA		Fair
	Maintenance	52	109	Venlafaxine 100-200	24 (22)	*P* < 0.001	
			116	Placebo	64 (55)		
Simon et al., 2004[127]	Acute	8	490	Venlafaxine 75-225	NA		Fair
	Continuation	26	161	Venlafaxine XR 75-225	45 (28)	*P* < 0.001	
			157	Placebo	82 (52)		

NA, not applicable; SR, slow release.
* Active treatment vs. placebo.

Nefazodone vs. placebo. Two trials, both rated fair, evaluated nefazodone.[125,132] In the relapse prevention study, investigators randomized patients in remission (HAM-D ≤ 10) to 36 weeks of double-blind treatment with nefazodone 400-600 mg/day (N = 65) or placebo (N = 66).[125] Statistically significantly fewer nefazodone-treated than placebo-treated patients relapsed (2 percent vs. 18 percent, respectively; $P = 0.009$). The recurrence prevention study openly treated 681 patients with chronic or recurrent major depression for 12 weeks with nefazodone 300-600 mg/day.[132] Patients who responded (≥ 50 percent improvement in HAM-D score from baseline) continued open-label nefazodone for an additional 16 weeks, and patients who maintained a response after this 16 weeks of continuation treatment were randomly assigned to 1 year of double-blind treatment with nefazodone (N = 76) or placebo (N = 84). The rate of recurrence was statistically significantly lower for patients on nefazodone than for those on placebo (30 percent vs. 48 percent, respectively; $P = 0.043$).

Paroxetine vs. placebo. One UK trial[119] and one US trial[133] assessed long-term treatment with paroxetine. Both trials randomized patients who had responded to acute-phase paroxetine therapy to 1 year of paroxetine or placebo. The UK study assessed relapse prevention after 16 weeks of double-blind treatment and recurrence prevention after an additional 36 weeks of continued double-blind treatment.[119] After 16 weeks, significantly fewer paroxetine-treated patients had relapsed than placebo-treated patients (3 percent vs. 19 percent, respectively; $P < 0.01$). Of the patients who maintained a response through the continuation phase and entered the maintenance phase, recurrence rates were lower for paroxetine-treated patients than for placebo-treated patients (14 percent vs. 30 percent, respectively; $P < 0.05$).

The US study was an extension of a 6-week acute-phase trial that compared paroxetine, imipramine, and placebo.[133] Investigators invited patients who had responded in the 6-week trial to continue flexible-dose, double-blind treatment for up to 1 year. Treatment allocation in the long-term extension was not randomized; the authors reported only aggregated relapse rates. More placebo-treated patients withdrew from the long-term trial because of "lack of efficacy"[133, page 25S] (n = 10; 22 percent) than did patients treated with either paroxetine (n = 11; 12 percent) or imipramine (n = 3; 4 percent).

Sertraline vs. placebo. One study assessed relapse prevention;[126] three other studies assessed recurrence prevention.[134,136,137] In the relapse prevention study, 295 patients who had responded in the acute phase were randomized to 44 weeks of double-blind treatment with sertraline 50-200 mg/day (N = 185) or placebo (N = 110).[126] Statistically significantly fewer sertraline-treated patients than placebo-treated patients experienced a relapse (13 percent vs. 46 percent, respectively; $P < 0.001$).

The good-quality relapse/recurrence prevention trial addressed potential methodological biases by including patients with recurrent depression who had been successfully treated for at least 4 months with any antidepressant other than sertraline.[136] Treatment was substituted with placebo for 2 months to identify patients truly in remission; patients who continued to remain in remission were randomized to sertraline 50 mg/day (N = 95), sertraline 100 mg/day (N = 94), or placebo (N = 99) and followed for 18 months. Patients treated with sertraline were statistically significantly less likely to have a recurrent depressive episode than patients treated with placebo (17 percent vs. 33 percent, respectively, for the pooled comparison; $P = 0.002$).

Similarly, the other two recurrence prevention studies found that patients treated with sertraline had fewer recurrences than did those on placebo.[134,137] In a 76-week maintenance phase, 6 percent of sertraline-treated and 23 percent of placebo-treated patients had a recurrent depressive episode ($P = 0.002$).[134] Differences did not reach statistical significance in a 100-week maintenance treatment of community residents 65 years of age and older with major depression; 45 percent of sertraline-treated patients and 54 percent of placebo-treated patients had a recurrent episode ($P = 0.21$).[137]

Venlafaxine vs. placebo. Two trials studied venlafaxine.[127,138] The relapse prevention study openly treated 490 patients with major depression with venlafaxine XR 75-225 mg/day for 8 weeks.[127] Patients who responded (CGI-S ≤ 3 and HAM-D ≤ 10) were randomized to 26 weeks of double-blind treatment with venlafaxine (N = 161) or placebo (N = 157). Statistically significantly fewer venlafaxine-treated patients than placebo-treated patients experienced a relapse (28 percent vs. 52 percent, respectively; $P < 0.001$).

The recurrence prevention study openly treated 495 patients with recurrent major depression for 6 months with venlafaxine 100-200 mg/day.[138] After 6 months, those who had responded (HAM-D ≤ 12) were randomized to 12 months of venlafaxine (N = 109) or placebo (N = 116). The recurrence rate was statistically significantly lower for venlafaxine-treated patients than for placebo-treated patients (22 percent vs. 55 percent, respectively; $P < 0.001$).

Treating Treatment-Resistant Depression or Relapse or Recurrence: Key Points

One head-to-head trial indicated that venlafaxine and paroxetine differ in their efficacy for treating major depression that has not responded to previous antidepressants.[141] This trial followed patients who were resistant to at least two previous antidepressant treatments; it found statistically significantly higher response and remission rates with venlafaxine than with paroxetine. Two studies yielded evidence of the effectiveness of one second-generation antidepressant compared with another for the treatment of depressive disorders in patients who had not responded to initial treatment.[140 146] One trial of good quality indicates that the compared treatments do not differ in their effectiveness as second-line agents.[146] An open-label Spanish study contradicts this finding; it reported statistically significant differences in effectiveness between compared treatments.[140] The contradictions between the one good-quality study and the two fair-quality studies led us to rate the overall strength of this evidence as moderate.

Treating Treatment-Resistant Depression or Relapse or Recurrence: Detailed Analysis

Head-to-head evidence. Three studies assessed differences among alternative antidepressants in patients who had either not responded or could not tolerate an acute-phase treatment (Table 19).[140,141,146] They covered a range of antidepressants; the common element was venlafaxine.

Bupropion SR vs. sertraline vs. venlafaxine XR. One effectiveness trial rated good quality assessed differences in effectiveness in patients with MDD who had not gone into remission (Quick Inventory of Depressive Symptomatology – Clinician version [QIDS-C-16] ≤ 5) or could not tolerate citalopram during acute-phase treatment.[146] Participants eligible for second-step treatment had the option of switching to an alternative medication, cognitive behavioral therapy, or augmentation therapy. To mimic clinical practice, patients could opt to exclude certain second-step treatment options, and they were then randomized to an acceptable treatment option. The investigators compared only the treatments for which patients had accepted randomization. Of the 727 patients randomized to a second-generation antidepressant, 239 received bupropion SR 150-400 mg/day, 238 received sertraline 50-200 mg/day, and 250 received venlafaxine XR 37.5-375 mg/day. Doses were adjusted based on clinical judgment and side effect rating scales. Second-step treatment was continued for up to 14 weeks.

Table 19. Head-to-head studies of treatment-resistant and recurrent depression

Study	Duration (Weeks)	N	Comparison and Dose (mg/d)	Response N (%)		Remission N (%)		Quality Rating
Baldomero et al., 2005[140]	24 (open)	1,465	Conventional therapy (pooled)	1,034 (71)	*P* < 0.001	754 (52)	*P* < 0.001	Fair
		294	Citalopram 20-40	209 (71)	*P* = 0.024	153 (52)	*P* = 0.02	
		248	Fluoxetine 20-40	174 (70)	*P* = 0.012	128 (52)	*P* = 0.03	
		116	Mirtazapine 30-45	75 (65)	*P* = 0.004	52 (45)	*P* = 0.003	
		312	Paroxetine 20-40	226 (73)	*P* = 0.078	161 (52)	*P* = 0.015	
		279	Sertraline 50-150	197 (71)	*P* = 0.014	147 (53)	*P* = 0.04	
		1,632	Venlafaxine 75-225	1,262 (78)		963 (59)		
Poirier and Boyer, 1999[141]	4	62	Paroxetine 30-40	18 (36)	*P* = 0.07	11 (18)	*P* = 0.02	Fair
		61	Venlafaxine 200-300	27 (45)		22 (37)		
Rush et al., 2006[146]	14	239	Bupropion 150-400	62 (26)	*P* = NR (ns)	51 (21)	*P* = 0.16	Good
		238	Sertraline 50-200	63 (27)		42 (18)		
		250	Venlafaxine 37.5-375	62 (25)		62 (25)		

NR, not reported; ns, not statistically significant.

At endpoint, response and remission rates were not statistically significantly different among bupropion SR, sertraline, and venlafaxine XR. For response, the figures were 26 percent, 27 percent, and 28 percent, respectively; (*P* = NR [ns]); for remission, the figures were 21 percent, 18 percent, and 25 percent, respectively (*P* = 0.16). Treatments also differed only minimally with respect to tolerability and adverse events.

Although several comparative studies included patients who had relapsed or who were experiencing a recurrent depressive episode in acute-phase management studies, no study *specifically* randomized patients to one second-generation antidepressant or another upon relapse or recurrence. The good-quality trial provides the most direct evidence relative to the second part of KQ 2; 75 percent of patients in this trial had failed acute treatment of a recurrent depressive episode. Among all patients in this trial, the investigators found no differences among bupropion SR, sertraline, and venlafaxine XR as an alternative treatment.

Venlafaxine vs. paroxetine or numerous other antidepressants. One effectiveness trial randomized 3,502 patients with major depression, dysthymia, or minor depression who had shown inadequate response or intolerance to at least 4 weeks of previous antidepressant treatment with venlafaxine XR 75-225 mg/day or with some other conventional antidepressant therapy.[140] Conventional therapy selection was at the discretion of the treating psychiatrist; it included citalopram 20-40 mg/day (N = 333), fluoxetine 20-40 mg/day (N = 292), mirtazapine 30-45 mg/day (N = 133), paroxetine 20-40 mg/day (N = 361), sertraline 50-150 mg/day (N = 299), and other miscellaneous drug treatments (N = 254).

After 24 weeks of treatment, venlafaxine-treated patients had a statistically significantly better rates of response and remission than patients treated with conventional therapy. (For response, the figures were 78 percent vs. 71 percent, respectively; *P* < 0.001; for remission, the figures were 59 percent vs. 52 percent, respectively; *P* < 0.001.) Response and remission rates for venlafaxine XR were statistically significantly better than each of the individual drugs characterized as conventional therapy, except for paroxetine.

The response and remission rates in this study were much higher than in the good-quality effectiveness trial just described.[146] Although differences in measurement scales may partially

explain response rates, the reason that remission rates differed remains unclear because both trials used a HAM-D cutoff point ≤ 7 to classify persons in remission.

One efficacy trial assessed differences between paroxetine and venlafaxine in patients with major depression who either had not responded or could not tolerate at least two previous treatments for their current depressive episode.[141] Patients were to be no more than minimally improved (CGI-I ≥ 3) with their second treatment. The investigators enrolled 123 patients in the study—61 on venlafaxine 200-300 mg/day and 62 on paroxetine 30-40 mg/day—and followed them for 4 weeks. At endpoint, statistically significantly more venlafaxine-treated patients than paroxetine-treated patients were classified as having responded to treatment (≥ 50 percent improvement in HAM-D from baseline; 45 percent vs. 36 percent, respectively; $P = 0.07$) and being in remission (HAM-D < 10; 37 percent vs. 18 percent, respectively; $P = 0.02$). The incidence of adverse events was comparable between treatment groups.

Key Question 3: Efficacy or effectiveness for treating symptoms accompanying depression

All Symptoms: Overview

For this issue, we focus on the comparative benefit of medications for patients with depression and an accompanying symptom cluster. We identified studies addressing six symptom clusters: anxiety, insomnia, pain, psychomotor change (retardation or agitation), melancholia (a depressive subtype that is a severe form of MDD with characteristic somatic symptoms), and somatization (physical complaints that are manifestations of depression rather than of an underlying physical illness). This set does not represent a complete list of symptoms commonly accompanying depression. For example, we did not identify any studies addressing fatigue, loss of energy, or appetite change—some of the more common accompanying symptoms reported by depressed patients.[147]

For each symptom cluster, we arrange our summary by how the data address two subquestions:

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

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

We identified 28 relevant trials (29 articles) (Tables 19-23); one trial was reported in two separate articles.[91,148] Twenty-two studies were head-to-head trials; one assessed three symptom subgroups.[149] We identified 11 head-to-head trials (12 articles) on anxiety,[35,37,38,54,64,78,85,91,108,148-150] six on insomnia,[41,60,81,82,96,151] three on melancholia,[69,149,152] two on pain,[71,153] and one each on psychomotor changes[149] and somatization.[49] One trial addressing somatization was an open-label effectiveness trial.[49] It did not meet our eligibility criteria because of the lack of double-blinding; however, we report on its results because it was a well-conducted randomized controlled effectiveness trial and comprises the only evidence for somatization in depressed patients. The

remaining six studies were placebo-controlled trials. Three addressed pain,[154-156] two addressed anxiety,[157,158] and one addressed melancholia.[159]

All but two studies[38,64] either were funded by or involved authors funded by pharmaceutical companies.

We rated all studies as fair quality with three exceptions (noted below). The fair rating was nearly universally a result of inadequate description of randomization and allocation concealment. A second common weakness was failure to report attrition rates, which occurred in 26 percent of trials.[149,150,154,155,157-159]

We rated the quality of the one effectiveness trial as not applicable as it did not meet our initial selection criteria.[49] In addition, we have two poor-quality studies, one on melancholia[152]and the other on pain.[153] Both of the poor studies had high attrition; one had high differential attrition between treatment groups,[152] and the other had high overall attrition.[153] We comment on data from these two studies because the evidence base for pain and melancholia is otherwise very weak. Poor studies were included only if the available evidence was very limited. For any poor studies retained for use in this report, we required, at a minimum, that investigators had employed a randomization scheme and applied ITT analysis.

Detailed information on these studies can be found in Evidence Table 3 in Appendix D. Evidence Table 4 provides information on systematic reviews and meta-analyses related to treating depression and accompanying symptoms.

Anxiety: Key Points

For KQ 3a, on the treatment of depression in patients with accompanying anxiety symptoms, we identified six head-to-head trials[64,78,91,108,148-150] and one placebo-controlled trial.[157] For KQ 3b, treatment of accompanying anxiety symptoms in patients with MDD, we included 10 head-to-head trials[35,37,38,54,64,78,85,108,148,150] and two placebo-controlled trials.[157,158] Six of these trials, in seven articles, addressed both key questions.[64,78,91,108,148,150,157] Of these 13 trials, five compared various SSRIs with each other or placebo, six compared an SSRI with an SNRI or another second-generation drug, and one each compared an SSRI or another second-generation drug only with placebo (Table 20). We rated the strength of evidence for both of these questions as moderate.

Table 20. Studies of adults with major depressive disorders and accompanying anxiety

Study	Interventions	N	Results	Quality Rating
SSRIs vs. SSRIs				
Chouinard et al., 1999[35]	Fluoxetine Paroxetine	203	KQ 3b. Improvement in anxiety scores was similar for both treatment groups (*P* = NR)	Fair
Fava et al., 1998[37]	Fluoxetine Paroxetine Placebo	128	KQ 3b. Improvement in anxiety scores was similar for both treatment groups and placebo (*P* = NR)	Fair
Fava et al., 2000[150]	Fluoxetine Paroxetine Sertraline	128 (all with anxiety)	KQ 3a. Improvement in depression scores (*P* = 0.323), depression response rates (*P* = 0.405) and remission rates were similar for all groups (*P* = 0.588) KQ 3b. Improvement in anxiety scores were similar for all 3 treatment groups (*P* = 0.199)	Fair

Flament et al., 1999[149]	Fluoxetine Sertraline	286 overall; 131 with anxiety	KQ 3a. Improvement in depression scores and depression response rates were similar for both treatment groups (*P* = NR)	Fair
Gagiano, 1993[38]	Fluoxetine Paroxetine	90	KQ 3b. Improvement in anxiety scores was similar for both treatment groups (*P* = NR)	Fair
SSRI vs. SNRI or other second-generation antidepressant				
Baldwin et al.,1996[85]	Paroxetine Nefazodone	206	KQ 3b. Improvement in anxiety scores was similar for both treatment groups (95% CI for difference, -0.7-3.8)	Fair
DeNayer et al., 2002[64]	Fluoxetine Venlafaxine	146 (all with anxiety)	KQ 3a. Improvement in depression scores was greater and response rates were higher for venlafaxine compared with fluoxetine (*P* < 0.05) KQ 3b. Improvement in anxiety scores was greater for venlafaxine compared with fluoxetine (*P* = 0.0004)	Fair
Leinonen et al., 1999[54]	Citalopram Mirtazapine	270	KQ 3b. Improvement in anxiety scores was similar for both treatment groups (*P* = 0.75)	Fair
Rush et al., 2001[108]	Sertraline Bupropion SR	248 overall; top quartile of HAM-A score with anxiety (number not provided)	KQ 3a. Depression response and remission rates were similar for both treatment groups (*P* = NR) KQ 3b. Improvement in anxiety scores was similar for both treatment groups (*P* = NR)	Fair
Sir et al., 2005[78]	Sertraline Venlafaxine XR	163 overall; 120 with anxiety	KQ 3a. Improvement in depression scores (*P* = 0.70), depression response rates (*P* = 0.26), and remission rates (*P* = 0.44) were similar for both groups KQ 3b. Improvement in anxiety scores was similar for both treatment groups (*P* = 0.32)	Fair
Trivedi et al., 2001[148] and Rush et al., 2001[91]	Sertraline Bupropion SR Placebo	724 overall; top quartile of HAM-A score with anxiety (number not provided)	KQ 3a. Depression response and remission rates were similar for both active groups and placebo (*P* = NR) KQ 3b. Improvement in anxiety scores was similar for treatment groups (*P* > 0.41)	Fair
SSRI or SNRI vs. Placebo				
Joliat et al., 2004[157]	Fluoxetine (weekly vs. daily) Placebo	799 overall; 374 with anxiety	KQ 3a. Depression relapse rates were similar for both medication groups and appeared better than those for placebo, but no statistical comparisons were reported (*P* = NR) KQ 3b. Worsening of anxiety scores appeared better for medication groups than for placebo, but there were no statistical comparisons (*P* = NR)	Fair
Khan et al., 1998[158]	Venlafaxine (3 doses) Placebo	403 overall; 346 with anxiety	KQ 3b. Improvement in anxiety scores for all 3 venlafaxine groups was superior to placebo group (*P* < 0.05); improvement was similar for the 3 venlafaxine dose groups	Fair

NR, not reported; SR, slow release.

Efficacy: KQ 3a: Depressive episode in patients with anxiety. Overall, six head-to-head studies and one fair placebo-controlled study indicated that antidepressant medications do not differ in treatment efficacy for depressed patients with accompanying anxiety symptoms. All

seven trials analyzed a subgroup with identified high anxiety. However, only three used the same definition criteria (a HAM-D anxiety-somatization factor of 7 or more).[78,150,157]

The head-to-head trials compared SSRIs with each other,[149,150] with bupropion SR,[91,108,148] and with SNRIs.[64,78] Studies appeared to compare similar doses of antidepressant medications. Two studies comparing SSRIs (including fluoxetine, paroxetine, and sertraline) found no statistically significant differences in depressive improvement, response rates, or remission rates.[149,150] Two studies comparing bupropion SR and sertraline found no significant differences in response or remission rates.[91,148] Two studies comparing an SSRI and venlafaxine showed mixed results. One found a greater decrease in depressive severity and higher response rates with venlafaxine than with fluoxetine,[64] and one found no statistically significant difference in depressive severity change, response rates, or remission rates between venlafaxine XR and sertraline.[78]

One placebo-controlled trial of patients whose MDD had already responded to treatment tested the efficacy of fluoxetine against placebo in preventing a relapse of depression during continuation-phase treatment.[157] Fluoxetine appeared to be more efficacious than placebo in preventing a relapse of a depressive episode, but statitistical comparsions were not reported.

Efficacy: KQ 3b: Anxiety in depressed patients. Overall, results from 10 head-to-head studies and two placebo-controlled studies suggested that antidepressant medications do not differ in treatment efficacy for treating anxiety associated with MDD. Seven of these 12 trials analyzed a subgroup with high anxiety.[64,78,108,148,150,157,158] Only three used identical definitions to identify the high anxiety group.[78,150,157] In addition, outcome definitions for anxiety varied. The studies compared similar doses of antidepressants.

The head-to-head trials compared SSRIs with each other, with SNRIs, and with other second-generation drugs (bupropion, nefazodone). Four studies comparing SSRIs (including fluoxetine, sertraline, and paroxetine) found no statistically significant differences for treatment of patients' anxiety symptoms.[35,37,38,150] Two studies comparing sertraline and bupropion SR found no difference in anxiety reduction.[108,148] Two studies comparing an SSRI (fluoxetine, sertraline) with venlafaxine found mixed results; venlafaxine produced a greater decrease in anxiety severity than fluoxetine in one study,[64] whereas the other study reported similar anxiety reduction for venlafaxine XR and sertraline.[78] One study comparing paroxetine and nefazodone found no difference in anxiety reduction.[85] A final head-to-head study comparing citalopram to mirtazapine found no difference in anxiety reduction.[54]

One placebo-controlled study comparing fluoxetine with placebo during continuation-phase treatment reported that anxiety worsened to a lesser degree with fluoxetine treatment, but the authors gave no statistical information.[157] A second placebo-controlled trial reported that venlafaxine treatment produced a statistically greater reduction in anxiety scores than placebo.[158]

Effectiveness: KQ 3a and KQ 3b. We identified no effectiveness trial relating to treatment of depression with accompanying anxiety symptoms. We expect, however, that if any such study were to be done, it would be less likely than the efficacy trials to show any differences between medications for this population.

Anxiety: Detailed Analysis

Head-to-head evidence. We identified 11 head-to head trials comparing the efficacy of specific medications treating depressed patients with coexisting anxiety symptoms. One study addressed only improvement in depression as an outcome (i.e., KQ 3a).[149] Five studies addressed only improvement in anxiety as an outcome (i.e., KQ 3b).[35,37,38,54,85] The remaining six articles addressed both questions.

Fluoxetine vs. paroxetine. Two studies compared the efficacy of low-to-high doses of fluoxetine with similar doses of paroxetine for treatment of anxiety.[35,38] Neither study required high anxiety for inclusion in the analysis.

One trial compared fluoxetine (20-80 mg/day) and paroxetine (20-50 mg/day) in a 12-week trial involving 203 patients with severe MDD.[35] Improvements on multiple measures of anxiety did not substantially differ between the two treatment groups.

A second study compared fluoxetine (20-60 mg/day) and paroxetine (20-40 mg/day) over 6 weeks in 90 patients with severe MDD.[38] Mean baseline anxiety severity was comparable, with each group having a moderate to severe degree of anxiety. Improvements in HAM-A scores were similar between the two groups.

Fluoxetine vs. paroxetine vs. placebo. One study compared low-to-high doses of fluoxetine, low-to-medium doses of paroxetine, and placebo in a pooled analysis of two multicenter trials, which were each 12 weeks in duration.[37] Patients had MDD of at least moderate severity. The analysis pooled data from five sites (not all) in the two trials; no explanation was provided for the limited inclusion. The outcome addressed was the effect of medications on co-occurring anxiety symptoms. Inclusion in the analysis did not require a high anxiety score; baseline mean Covi Anxiety Scale scores were similar (< 7) in all groups, consistent with patients not being anxious. Improvement in anxiety symptoms on the Covi Anxiety Scale did not differ among the three groups.

Fluoxetine vs. paroxetine vs. sertraline. One RCT compared low-to-high dose fluoxetine (20-60 mg/day), low-to-high dose paroxetine (20-60 mg/day), and low-to-high dose sertraline (50-200 mg/day) over 10 to 16 weeks in patients with MDD of at least moderate severity and high anxiety (as defined by a score on the six-item HAM-D anxiety-somatization scale ≥ 7 [range 0-18]).[150] Analyses were performed in the subgroup with high anxiety (n = 108 patients from a trial with 284 participants overall); the outcomes included both depressive measures and anxiety measures. Depressive outcomes between the three medications were similar, as measured by improvement in HAM-D total scores, by response rates (≥ 50 percent reduction in HAM-D score; fluoxetine, 73 percent, paroxetine, 77 percent; and sertraline, 86 percent, $P = 0.405$), and by remission rates (HAM-D endpoint ≤ 7; fluoxetine, 53 percent; paroxetine, 50 percent; and sertraline, 62 percent; $P = 0.588$). Likewise, authors reported no difference between the three groups with respect to anxiety outcomes (measured by overall change on HAM-D anxiety-somatization score).

Fluoxetine vs. sertraline. One study compared low-to-medium doses of fluoxetine (20-40 mg/day) and sertraline (50-100 mg/day) over 6 weeks in patients with MDD of at least moderate severity who also had high anxiety as defined by a Covi Anxiety Scale score ≥ 7.[149] The outcome

was depression response. Authors reported that response rates (defined by ≥ 50 percent reduction in HAM-D total score) did not differ between the fluoxetine-treated group (48 percent) and the sertraline-treated group (47 percent).

Citalopram vs. mirtazapine. One study compared the efficacy of low-to-high dose citalopram (20-60 mg/day) and low-to-high dose mirtazapine (15-60 mg/day) over 8 weeks in 270 patients with MDD of at least moderate severity.[54] The outcome was treatment effect on anxiety as measured by HAM-A scores. However, patients were not categorized by anxiety level, and the analysis included all patients with MDD, not merely those with anxiety. The improvement in anxiety symptoms did not differ between citalopram and mirtazapine (mean HAM-A change in both groups was approximately -13 points).

Fluoxetine vs. venlafaxine. One trial compared low-to-medium doses of fluoxetine (20-40 mg/day) with low doses of venlafaxine (75-150 mg/day) over 12 weeks in 146 moderately depressed patients with MDD who had a Covi Anxiety Scale score ≥ 8 (consistent with clinically relevant anxiety).[64] Both depression and anxiety outcomes were reported. The improvement in depressive severity on the HAM-D was significantly greater in the venlafaxine-treated group than the fluoxetine-treated group (-14.4 points vs. -10.4 points, $P = 0.0048$). Similarly, the mean reduction on the Covi Anxiety Scale was greater for venlafaxine than for fluoxetine (-5.7 points vs. -3.9 points, $P = 0.0004$).

Sertraline vs. bupropion SR. One efficacy trial compared low-to-high dose sertraline with low-dose Bupropion SR over 16 weeks in 248 patients with MDD of moderate severity.[108] High anxiety patients were defined as those with scores in the top quartile on HAM-A (≥ 19, consistent with at least moderate anxiety). Outcomes included both depression (HAM-D-21) and anxiety (HAM-A) measures. For the subgroup with high anxiety, depression response rates (≥ 50 percent reduction in total score, approximately 70 percent in each group) and remission rates (endpoint ≤ 8, approximately 70 percent in each group) were similar. Likewise, in the high-anxiety subgroup, authors reported no difference in anxiety reduction (measured by mean change in HAM-A) between patients treated with sertraline (-10.0) and bupropion (-9.7).

Sertraline vs. bupropion SR vs. placebo. One pooled analysis of two 8-week RCTs compared the efficacy of low-to-high dose sertraline (50-200 mg/day), low-to-high dose Bupropion SR (150 mg-400 mg/day), and placebo in 724 patients with MDD of at least moderate severity.[91,148] One set of investigators reported on depressive outcomes;[91] the other set reported on anxiety outcomes.[148]

The two sets of investigators defined high anxiety in slightly different ways. In the study on depressive outcomes,[91] the high anxiety subgroup comprised patients with a HAM-A score in the top quartile of enrolled patients (HAM-A ≥ 25). For this subgroup, rates of depression response (defined as HAM-D-21 reduction of ≥ 50 percent; estimated by us from the figure in the article to be approximately 60 percent to 70 percent for each of three arms) and of remission (HAM-D-21 ≤ 8; estimated to be 25 percent to 35 percent for each arm) did not differ by treatment group. Furthermore, the authors did not report any significant relationship between quartile of baseline anxiety and antidepressant response for any of the three treatment arms.

In the study on anxiety outcomes,[148] investigators defined baseline anxiety as minimal (HAM-A ≤ 14), moderate (HAM-A 15-19), or severe (HAM-A ≥ 20). They reported no

differences in mean HAM-A reduction between patients treated with bupropion SR and those treated with sertraline for any of the anxiety severity subgroups (severe, bupropion SR ~ -15 points vs. sertraline ~ -13 points; moderate, bupropion SR ~ -8 vs. sertraline ~ -9; minimal, bupropion SR ~ -6 vs. sertraline ~ -5; all data estimated from figures in the paper).

Sertraline vs. venlafaxine XR. One efficacy study compared low-to-high dose sertraline (50-150 mg/day) with low-to-high dose venlafaxine XR (75-225 mg/day) over 8 weeks in a subgroup of 120 patients with MDD of at least moderate severity and accompanying anxiety, defined as a HAM-D anxiety-somatization score of ≥ 7.[78] Outcomes included both depressive (HAM-D-17) and anxiety (HAM-A) measures. Authors reported no difference between treatment groups in mean depressive severity reduction (-17.3 for sertraline vs. -14.8 for venlafaxine XR, $P = 0.7$), depression response rates (≥ 50 percent reduction in total score, 79.6 percent for sertraline vs. 68.9 percent for venlafaxine XR, $P = 0.26$), or depression remission rates (endpoint ≤ 7, 63.0 percent for sertraline vs. 54.1 percent with venlafaxine XR, $P = 0.44$).

Anxiety symptom outcomes did not differ between treatment groups for the overall study population (N = 163) or for the high anxiety subgroup (n = 120). In the overall study population, the mean reduction in HAM-A was -14.1 for the sertraline-treated group and -12.9 for the venlafaxine XR-treated group ($P = 0.32$). In the high anxiety subgroup, response on the HAM-D anxiety-somatization subscale (criteria not described) was similar for both treatment arms (83.3 percent for sertraline vs. 70.5 percent for venlafaxine XR, $P = 0.12$).

Paroxetine vs. nefazodone. One RCT compared the low-to-medium dose paroxetine (20-40 mg/day) with low-to-high dose nefazodone (200-600 mg/day) for treatment of accompanying anxiety symptoms over 8 weeks in patients with moderate to severe MDD.[85] Inclusion in the analysis did not require high anxiety, and patients were not categorized based on anxiety level; the outcome was the mean difference between treatment groups in HAM-A improvement. Authors reported similar improvement in HAM-A for the treatment groups (-8.0 for paroxetine vs. -6.5 for nefazodone, P = NS, 95% CI for difference between groups, -0.7-3.8).

Placebo-controlled evidence. Two trials compared efficacy of a second-generation antidepressant only against placebo. One involved an SSRI (for both KQ 3a and 3b), the other examined an SNRI (for only KQ 3b).

Fluoxetine vs. placebo. One study compared the efficacy of two different preparations of fluoxetine and placebo in preventing depression relapse in patients whose depression had been successfully treated.[157] The study involved continuation-phase treatment, where the clinical goal was to prevent relapse of a successfully treated depressive episode.

The authors pooled data from two 25-week RCTs. Patients who were in remission (from study 1) or who responded (from study 2) to approximately 3 months of open-label fluoxetine treatment were randomly assigned to placebo, continued treatment with 20 mg/day fluoxetine, or (in study 2 only) 90 mg/week delayed-release fluoxetine. High anxiety patients were defined as those with a HAM-D-17 anxiety-somatization subscale score of ≥ 7 at baseline.

In the high anxiety subgroup (n = 374), depression relapse rates appeared to be lower in the fluoxetine daily and fluoxetine weekly groups (27.8 percent and 28.5 percent, respectively) than in the placebo group (53.3 percent); the authors did not provide statistical information.

Anxiety levels increased (worsened) for all treatment arms in the high anxiety subgroup. This increase appeared less in the fluoxetine daily and weekly groups (1.92 and 1.93) than in the placebo group (3.12), but again statistical significance was not reported.

Venlafaxine vs. placebo. One 12-week study randomly assigned patients with severe MDD to one of three doses of immediate-release venlafaxine—75 mg/day (low), 150 mg/day (low), or 200 mg/day (medium)—or to placebo.[158] Inclusion did not require a high anxiety score. Treatment effects on anxiety were analyzed in a subgroup of 346 patients with accompanying anxiety (defined as a score of ≥ 2 [at least moderate] on the HAM-D-17 anxiety-psychic item, range 0-4). Each treatment arm had an equivalent number of patients with high anxiety. All four treatment arms experienced a reduction in anxiety. Patients in all three venlafaxine groups had statistically significant greater improvement in HAM-D anxiety-psychic and anxiety-somatization scores compared with the placebo group. The three venlafaxine groups did not differ from each other in anxiety outcomes.

Insomnia: Key Points

We identified six head-to-head studies that compared the effects of medications on treatment of depression and accompanying insomnia (Table 21).[41,60,81,82,96,151] Three of these trials required insomnia for inclusion in the analysis, although the definitions used varied by study and consisted of brief 1- or 3-item measures.[41,81,151] Three other trials did not require insomnia for inclusion but rather assessed sleep for all subjects.[60,82,96] The three studies that identified an insomnia group provided data addressing both KQ 3a (effects on depressive symptoms) and KQ 3b (effects on insomnia).[41,81,151] The other studies provided information solely on insomnia outcomes. We rated the strength of evidence for both depression outcomes in patients with accompanying insomnia (KQ 3a) and insomnia outcomes in patients with depression (KQ 3b) as low. All studies were of fair quality.

Table 21. Studies of adults with major depressive disorders and accompanying insomnia

Study	Interventions	N	Results	Quality Rating
Fava et al., 2002[41]	Fluoxetine Paroxetine Sertraline	284 overall; 125 with insomnia	KQ 3a. Improvement in depression scores was similar for all groups (P = 0.853) KQ 3b. Improvement in sleep was similar for all groups (P = 0.852)	Fair
Lader et al., 2005[151]	Citalopram Escitalopram Placebo	1,321 overall; 638 with insomnia	KQ 3a. Improvement in depression scores for escitalopram was superior to citalopram and placebo (P < 0.05) KQ 3b. Improvement in sleep for escitalopram was superior to citalopram and placebo (P < 0.01)	Fair
Rush et al., 1998[81]	Fluoxetine Nefazodone	125 (all with insomnia)	KQ 3a. Improvement in depression scores (95% CI for difference between groups, -1.7-2.8) and depression response rates (P = NR) were similar for both groups KQ 3b. Improvement in sleep for nefazodone was superior to fluoxetine (P < 0.05)	Fair
Beasley et al., 1991[82]	Fluoxetine Trazodone	126	KQ 3b. Improvement in sleep scores was greater for trazodone than for fluoxetine (P = 0.001)	Fair
Cunningham et al., 1994[96]	Venlafaxine Trazodone	227	KQ 3b. Improvement in sleep scores was greater for trazodone than venlafaxine (P < 0.05)	Fair
Versiani et al., 2005[60]	Fluoxetine Mirtazapine	299	KQ 3b. Sleep quality improved similarly for both groups (overall score not reported)	Fair

Efficacy: KQ 3a: Depressive episode in patients with insomnia. Three head-to-head studies provide mixed evidence regarding comparative efficacy of medications for treatment of depression in patients with accompanying insomnia.[41,81,151] One study found a slightly greater reduction in depressive severity for escitalopram than for citalopram or placebo.[151] The other two studies showed no statistically significant differences in depressive outcomes for fluoxetine compared with paroxetine and sertraline[41] and for fluoxetine compared with nefazodone.[81]

Efficacy: KQ 3b: Insomnia in depressed patients. Six head-to-head studies provided mixed evidence about the effects of antidepressants on insomnia in patients with depression. Two studies reported greater improvement in sleep scores for trazodone than for fluoxetine[82] and venlafaxine.[96] However, neither of these trials analyzed a subgroup of patients with insomnia. One study each reported better sleep outcomes for escitalopram than for citalopram[151] and for nefazodone than for fluoxetine.[81] The clinical meaning of the small sleep outcome differences reported in these studies is unclear. One study each found no statistically significant differences between fluoxetine, paroxetine, and sertaline[41] and between fluoxetine and mirtazapine.[60]

Effectiveness. We identified no effectiveness studies concerning depression and insomnia.

Insomnia: Detailed Analysis

Head-to-head evidence. Six head-to-head trials addressed this issue.

Citalopram vs. escitalopram vs. placebo. One study compared low-to-medium dose citalopram (20-40 mg/day) with low-to-medium dose escitalopram (10-20 mg/day) in a pooled secondary analysis of three 8-week RCTs of patients with MDD of at least moderate severity as measured by the MADRS (10 items, individual score range 0-6, total score range 0-60).[151] Insomnia was defined as a score of 4 or greater on the single MADRS sleep item (range 0-6). Among 638 patients meeting insomnia criteria, depressive symptoms improved (i.e., MADRS scores declined) for all three treatment arms. Improvement was greater for escitalopram than for citalopram and placebo (escitalopram, -16.47; citalopram, -14.02; placebo, -12.2; $P < 0.01$ for escitalopram vs. citalopram, $P < 0.001$ for escitalopram vs. placebo; P = NR, not significant for citalopram vs. placebo).

Insomnia results also favored escitalopram. Mean improvement on the MADRS sleep item was better in the escitalopram group than in the citalopram and placebo groups (escitalopram, -1.65; citalopram, -1.31; placebo, -1.26; $P < 0.01$ for escitalopram vs. citalopram, $P < 0.01$ for escitalopram vs. placebo, P = NS for citalopram vs. placebo). Escitalopram-treated patients were more likely than others to achieve improvement in insomnia, defined as a score of 0 or 1 on the MADRS sleep item at week 8 (43.6 percent for escitalopram, 28.4 percent for citalopram, 24.4 percent for placebo, overall $P < 0.001$).

Fluoxetine vs. paroxetine vs. sertraline.One study compared low-to-high doses of fluoxetine (20-60 mg/day), paroxetine (20-60 mg/day), and sertraline (50-200 mg/day) in a study of MDD patients with at least a moderate degree of depression that lasted between 10 and 16 weeks.[41] A secondary analysis evaluated depression outcomes in patients with insomnia, defined as a score of at least 4 points on the HAM-D sleep disturbance subscale (a 0 to 6 scale consisting of a summed score of three HAM-D-17 sleep items [assessing initial, middle, and terminal insomnia],

where higher scores indicated worse insomnia). For the 125 patients in this subgroup, the three SSRIs did not differ significantly on the HAM-D score (overall $P = 0.853$).

This study also assessed the effect of medications on insomnia. Again, treatment groups did not differ. Insomnia (measured as above on the 6-point scale) improved to a similar degree for all three groups (fluoxetine, -3.1; paroxetine, -2.9; sertraline, -3.1; overall $P = 0.852$).

Fluoxetine vs. trazodone. One study compared low-dose fluoxetine (95 percent of participants took 20 mg/day) with low-to-medium dose trazodone (50-400 mg/day, median 250 mg) over 6 weeks in patients with major depression.[82] Investigators did not require insomnia symptoms for inclusion and did not analyze an insomnia subgroup. Overall HAM-D sleep disturbance scores improved more in the trazodone group than in the fluoxetine group (-2.7 vs. -1.6; $P = 0.001$).

Fluoxetine vs. nefazodone. One study compared low-to-moderate dose fluoxetine (20-40 mg/day) and nefazodone (200-500 mg/day) in a pooled analysis of three 8-week RCTs of patients with MDD of at least moderate severity, as measured by HAM-D-17.[81] The analysis was conducted on a subgroup of 122 patients with insomnia, defined by patient self-report at study entry of difficulty falling asleep each night, waking up during the night, or being unable to fall asleep again after getting out of bed. Depressive outcomes did not differ between treatment groups. The mean improvement in HAM-D was 12.2 points for fluoxetine and 11.4 for nefazodone (95% CI for difference, -1.7-2.8). Response rates, defined as HAM-D < 10, were essentially identical (45 percent for fluoxetine, 47 percent for nefazodone).

Sleep outcomes from the same study favored nefazodone. Patients receiving nefazodone had a mean improvement of -2.3 points on the HAM-D sleep disturbance items (range 0-6); the improvement for patients receiving fluoxetine was -1.6 ($P < 0.05$). Nefazodone-treated patients also had greater improvement on a secondary sleep measure, the Inventory for Depressive Symptomatology sleep items relating to early, middle, and late insomnia and hypersomnia (range 0-12, scored such that higher scores are better); patients receiving nefazodone improved by 2.4 points on this measure, compared with 1.7 points for patients treated with fluoxetine ($P < 0.01$).

Fluoxetine vs. mirtazapine. One study compared low-to-medium doses of fluoxetine (20-40 mg/day) with low-to-high doses of mirtazapine in an 8-week study of patients with severe MDD.[60] The investigators did not categorize subgroups of patients by the presence or absence of insomnia. They compared outcomes on the Leeds Sleep Evaluation Questionnaire for all trial participants. Total scores were not reported; efficacy on individual items did not differ in any substantial or consistent way between treatment groups.

Venlafaxine vs. trazodone vs. placebo. One study compared low-to-medium doses of venlafaxine (75-200 mg/day) and trazodone (150-400 mg/day) over 6 weeks in patients with major depression.[96] Investigators did not require insomnia symptoms for inclusion and did not analyze an insomnia subgroup. HAM-D sleep disturbance scores were better (lower) at endpoint in patients receiving trazodone than in those receiving either venlafaxine or placebo (score 1.42 for trazodone, 2.22 for venlafaxine, 1.95 for placebo; $P < 0.05$). HAM-D sleep disturbance factor scores at endpoint did not differ between venlafaxine and placebo (P = NR).

Placebo-controlled evidence. We identified no placebo-controlled trials for depression and insomnia.

Melancholia: Key Points

We identified three head-to-head studies[69,149,152] and one placebo-controlled study[159] (Table 22). All addressed KQ 3a: whether, for patients with melancholia, medications differed in their effect on depressive symptoms. All but one study was rated fair quality; one was rated poor. We rated the strength of evidence for the comparative efficacy and effectiveness of second-generation antidepressants for treating depression in patients with melancholia as low.

Table 22. Studies of adults with major depressive disorders and accompanying melancholia

Study	Interventions	N	Results	Quality Rating
Flament et al., 1999[149]	Fluoxetine Sertraline	286 overall; 197 with melancholia	KQ 3a. Depression response rates for sertraline were superior to fluoxetine ($P < 0.05$); improvement in depression scores was similar for both groups (P = NR)	Fair
Clerc et al., 1994[152]	Fluoxetine Venlafaxine	68 (all with melancholia)	KQ 3a. Improvement in depression scores was better for venlafaxine than fluoxetine ($P = 0.027$); response rates did not differ ($P = 0.08$)	Poor
Tzanakaki et al., 2000[69]	Fluoxetine Venlafaxine	109 (all with melancholia)	KQ 3a. Depression response and remission rates were similar for both groups (P = NR)	Fair
Mallinckrodt et al., 2005[159]	Duloxetine Placebo	2,342 overall; 1,572 with melancholia	KQ 3a. Improvement in depression scores was better for duloxetine than placebo ($P < 0.001$)	Fair

NR, not reported.

We found no evidence addressing the comparative efficacy and effectiveness of second-generation antidepressants for the treatment of accompanying melancholic symptoms (KQ 3b).

Efficacy: KQ 3a: Depressive episode in patients with melancholia. Three head-to-head studies compared fluoxetine with sertraline[149] or venlafaxine.[69,152] One study found a greater response rate in patients receiving sertraline than fluoxetine.[149] Another found no difference between the fluoxetine and venlafaxine groups in response and remission rates.[69] One poor-quality study found a greater decrease in depressive severity for venlafaxine than for fluoxetine but only a nonsignificant tendency toward a greater rate of response (a more robust outcome).[152]

One placebo-controlled study, a pooled data analysis of duloxetine trials, found a greater decrease in depressive severity for duloxetine than for placebo.[159]

Efficacy: KQ 3b: Melancholia in depressed patients. We identified no efficacy trials addressing treatment of melancholic symptoms.

Effectiveness: KQ 3a and KQ 3b. We identified no effectiveness trials for depressed patients with accompanying melancholia.

Melancholia: Detailed Analysis

Head-to-head evidence. We identified two fair-quality studies[69,149] and one poor-quality study,[152] all 6 weeks in length.

Fluoxetine vs. sertraline. One study enrolled patients who were at least moderately depressed (either MDD or the depressed phase of bipolar disorder); patients were randomized to low-to-medium dose fluoxetine (20-40 mg/day) or sertraline (50-100 mg/day) for 6 weeks.[149] In the subgroup with melancholia by DSM-III-R criteria, depression response rates (≥ 50 percent decrease in HAM-D-17) were significantly better for sertraline than for fluoxetine (59 percent vs. 44 percent, $P < 0.05$).

Fluoxetine vs. venlafaxine. Two studies provided mixed results on the relative efficacy of moderate-to-high doses of fluoxetine and venlafaxine. One trial involved severely depressed hospitalized patients or outpatients with MDD and melancholia per DSM-IV criteria; patients were randomized to either 60 mg/day of fluoxetine or 225 mg/day of venlafaxine.[69] Authors reported no statistically significant difference in response rates (≥ 50 percent decrease in HAM-D-21 or MADRS and CGI improvement score of 1 or 2) between groups (58 percent for fluoxetine, 65 percent for venlafaxine). Similarly, remission rates (final HAM-D-21 score < 7) did not differ significantly (fluoxetine, 35.8 percent; venlafaxine, 40.7 percent).

The other study (rated poor) involved severely depressed hospitalized patients with MDD and melancholia per DSM-III-R criteria; investigators randomized patients to either 200 mg/day of venlafaxine or 40 mg/day of fluoxetine.[152] Using ITT approaches, the investigators determined that the improvement in depressive severity (mean decrease in HAM-D-21 score) was significantly greater in the venlafaxine group than in the fluoxetine group (-18 points vs. -12.4 points, $P = 0.027$). Response rates (≥ 50 percent decrease in HAM-D severity) did not differ significantly between groups (73 percent venlafaxine vs. 50 percent fluoxetine; $P = 0.08$).

Placebo-controlled evidence. ***Duloxetine vs. placebo.*** One study compared low-to-high doses of duloxetine (40-120 mg/day) with placebo for MDD patients with at least moderate depressive severity.[159] The authors pooled results of eight RCTs of 9 weeks' duration (all part of the New Drug Application to the FDA for duloxetine) and identified a subgroup of patients (n = 1,572) with melancholia, per DSM-IV criteria, on whom to conduct this secondary analysis. Accordingly, the randomization was not stratified by a melancholic designation. Mean reductions in HAM-D-17 score were 8.97 for patients receiving duloxetine and 6.57 for those receiving placebo ($P < 0.001$), suggesting a benefit for duloxetine, on average, of slightly more than 2 points.

Pain: Key Points

We included two head-to-head trials[71,153] and three placebo-controlled trials[154-156] that assessed the efficacy of antidepressants for treatment of depression and accompanying pain symptoms (Table 23). One placebo-controlled trial required baseline pain for inclusion; this study provided data addressing both KQ 3a (depression outcomes in patients with accompanying pain) and KQ 3b (pain outcomes in MDD patients).[154] The other four trials did not require pain for inclusion, but rather assessed pain symptoms for all subjects; these trials provided

information for KQ 3b (pain outcomes) only.[71,153,155,156] All but one study was rated fair quality; one was rated poor. We rated the strength of evidence for both questions as low.

Table 23. Studies of adults with major depressive disorders and accompanying pain

Study	Interventions	N	Results	Quality Rating
Brannan et al., 2005[154]	Duloxetine Placebo	282	KQ 3a. Improvement in depression scores (*P* = 0.544), depression response rates (*P* = 0.901), and remission rates (*P* = 0.887) was similar KQ 3b. Improvement in pain scores was similar (*P* = 0.066)	Fair
Goldstein et al., 2004[153]	Duloxetine Paroxetine Placebo	353	KQ 3b. Improvement in pain scores was similar between active medications (*P* = NR), between paroxetine and placebo (*P* = 0.088), and between duloxetine 40 mg and placebo (*P* = 0.172); improvement in pain for duloxetine 80 mg was superior to placebo (*P* = 0.005).	Poor
Detke et al., 2002[155]	Duloxetine Placebo	245	KQ 3b. Pain score improvement was slightly greater for duloxetine than placebo (*P* = 0.019)	Fair
Detke et al., 2002[156]	Duloxetine Placebo	267	KQ 3b. Pain score improvement was slightly greater for duloxetine than placebo (*P* = 0.037)	Fair
Detke et al., 2004[71]	Duloxetine Paroxetine Placebo	367	KQ 3b. Improvement in pain scores was similar between duloxetine 80 mg and placebo (*P* = 0.063), and between duloxetine 120 mg and placebo (*P* = 0.086); improvement in pain for paroxetine was superior to placebo (*P* = 0.035)	Fair

NR, not reported.

Efficacy: KQ 3a: Depressive episode in patients with pain. One study found no difference in efficacy between duloxetine and placebo for treatment of depression in patients with mild to moderate pain.[154]

Efficacy: KQ 3b: Pain in depressed patients. One fair-quality trial and one poor-quality trial reported similar efficacy for duloxetine and paroxetine for treating pain symptoms in MDD patients.[71,153] Five placebo-controlled studies provided mixed evidence for efficacy of active drugs compared to placebo for treatment of accompanying pain. Five trials compared duloxetine with placebo;[71,153-156] three of these reported statistically greater pain improvement in at least one duloxetine treatment arm.[153,155,156] Two studies compared paroxetine with placebo;[71,153] one found a statistically greater improvement for paroxetine.[71] Overall, mean differences in pain scores between groups were small and may not be clinically meaningful.

No studies evaluated the efficacy of antidepressants in a subgroup of patients with moderate to severe pain. For outcome measures, all five studies used a visual analog scale (VAS) for overall pain (0 mm to 100 mm scale, where higher scores indicate worse pain); one trial also used the Brief Pain Inventory (BPI) severity scale (0 to 10 scale, where higher scores indicate worse pain).[154] No study reported percentages of patients with clinically important improvement in pain. All studies were funded by the maker of duloxetine.

Effectiveness: KQ 3a: Depressive episode and KQ 3b: Accompanying symptoms. We identified no effectiveness trials for treatment of patients with depression and accompanying pain.

Pain: Detailed Analysis

Head-to-head evidence. ***Duloxetine vs. paroxetine vs. placebo.*** Two multicenter trials compared the efficacy of duloxetine, paroxetine, and placebo. Pain symptoms were not required for inclusion in either study. Moreover, baseline pain severity was mild in both trials.

One trial compared two high doses of duloxetine (80 mg/day and 120 mg/day) to low-dose paroxetine (20 mg/day) and placebo.[71] Improvement in overall pain (decrease in 100 mm VAS) was similar for both duloxetine formulations and paroxetine (duloxetine 80 mg/day, -11.2 mm; duloxetine 120 mg/day, -12.2 mm; paroxetine, -16.0 mm; $P = 0.77$ for duloxetine 80 mg vs. paroxetine; $P = 0.66$ for duloxetine 120 mg vs. paroxetine). Mean pain improvement was statistically significantly superior to placebo for paroxetine ($P = 0.035$) but not for either duloxetine formulation ($P = 0.063$ for duloxetine 80 mg vs. placebo; $P = 0.086$ for duloxetine 120 mg vs. placebo).

One multicenter trial (rated poor quality) compared the efficacy of low-dose duloxetine (40 mg/day), high-dose duloxetine (80 mg/day), and low-dose paroxetine (20 mg/day) for treatment of accompanying pain in patients with MDD over 8 weeks.[153] Statistical comparisons between active treatment groups were not reported, but median change from baseline to endpoint in VAS overall pain was minimal and similar for all groups (-4 mm for duloxetine 40 mg, -7.5 mm for duloxetine 80 mg, -3.0 mm for paroxetine, 0 mm for placebo). Pain improvement in the duloxetine 80 mg group was small but statistically significantly better than placebo ($P = 0.005$). Median improvement in pain scores did not differ between duloxetine 40 mg and placebo ($P = 0.172$) or between paroxetine and placebo ($P = 0.088$).

Placebo-controlled evidence. ***Duloxetine vs. placebo*****.** One multicenter trial compared high-dose duloxetine (60 mg/day) with placebo over 7 weeks for treating patients with MDD and pain symptoms.[154] Participants were 282 outpatients who met DSM-IV criteria for major depression and reported accompanying pain, with a BPI average pain score of 2 or more at baseline. Patients who had "a primary pain complaint with a diagnosis such as arthritis, fibromyalgia, migraine headache or acute injury" were excluded. Mean baseline pain severity was moderate (BPI average: 4.85 for duloxetine, 4.62 for placebo). The authors found no statistically significant difference between duloxetine and placebo on either depression or pain outcomes. Mean HAM-D-17 improvement was similar for the groups (duloxetine, -10.9; placebo, -10.3; $P = 0.544$). Depression response and remission rates did not differ between duloxetine and placebo (response 42 percent vs. 40 percent, $P = 0.901$; remission 23 percent vs. 24 percent, $P = 0.887$). Mean reduction in BPI average pain was similar for duloxetine and placebo (-2.32 vs. -1.80; $P = 0.066$). Mean changes in BPI worst pain, least pain, and current pain intensity did not differ between treatment groups ($P > 0.10$ for all comparisons). Mean changes in VAS overall pain did not differ between treatment groups (values NR, P = NR).

Two trials compared the efficacy of high-dose duloxetine (60 mg/day) to placebo over 9 weeks for treatment of pain in patients with depression.[155,156] Inclusion criteria were similar in both studies: participants met DSM-IV criteria for MDD but were not required to have pain symptoms. Mean baseline pain severity was mild (VAS for overall pain: 29.0 and 25.4 for duloxetine, 28.2 and 26.2 for placebo). Both studies reported small but statistically significant differences in VAS overall pain improvement favoring duloxetine over placebo: -8.5 mm vs. -1.3 mm ($P = 0.019$)[155] and -11.0 mm vs. -6.4 mm ($P = 0.037$).[156]

Psychomotor Change: Key Points

One head-to-head trial addressed KQ 3a on depression response in subgroups with psychomotor retardation or psychomotor agitation (Table 24).[149] We rated the strength of evidence for this issue as low. We found no evidence for the comparative efficacy and effectiveness of second-generation antidepressants for the treatment of accompanying psychomotor symptoms (KQ 3b).

Table 24. Studies of adults with major depressive disorders and accompanying psychomotor change

Study	Interventions	N	Results	Quality Rating
Flament et al., 1999[149]	Fluoxetine Sertraline	286	KQ 3a. In patients with psychomotor retardation, depression scores and response rates were similar for both groups (*P* = NR) In patients with psychomotor agitation, depression scores (*P* = 0.02) and response rates (*P* = 0.04) were superior for sertraline	Fair

NR, not reported.

Efficacy: KQ 3a: Depressive episode in patients with psychomotor changes. One study provided evidence that fluoxetine and sertraline have similar efficacy for treatment of depression in patients with psychomotor retardation. The same study reported that sertraline was more efficacious than fluoxetine for treatment of depression in patients with psychomotor agitation.[149]

Efficacy: KQ 3b: Psychomotor changes in depressed patients. We identified no efficacy trials addressing treatment of psychomotor change symptoms.

Effectiveness: KQ 3a: Depressive episode and KQ 3b: Accompanying symptoms. We found no effectiveness trials concerning patients with depression and accompanying psychomotor problems.

Psychomotor Change: Detailed Analysis

Head-to-head evidence. ***Fluoxetine vs. sertraline.*** One 6-week trial compared low-to-medium doses of fluoxetine and sertraline for treating depression in subgroups of patients with MDD or the depressed phase of bipolar disorder and psychomotor retardation or psychomotor agitation.[149] The subgroup with psychomotor retardation comprised 47 patients with a score of 2 or more on HAM-D item 8 (retardation) and 1 or less on item 9 (agitation). In this subgroup, mean HAM-D scores improved similarly for fluoxetine- and sertraline-treated patients (-10.7 vs. -9.1 points, *P* = NR). Response rates (≥ 50 percent improvement on HAM-D-17 total score) were also similar for fluoxetine and sertraline (46 percent vs. 48 percent, *P* = NR). The same study evaluated depression response in a subgroup of 78 patients with psychomotor agitation, defined as a score of 1 or less on HAM-D item 8 and 2 or more on item 9. Among patients with psychomotor agitation, improvement in HAM-D total score was greater in patients receiving sertraline than in those receiving fluoxetine (-12.4 vs. -8.7 points, *P* = 0.02). Response rates were also significantly better for sertraline than for fluoxetine (62 percent vs. 39 percent, *P* = 0.04).

Placebo-controlled evidence. We identified no placebo-controlled trials involving this patient population.

Somatization: Key Points

We identified one open-label, head-to-head effectiveness trial that compared effects of medications on accompanying somatization in depressed primary-care patients (KQ 3b) (Table 25).[49] The grade of evidence for the comparative efficacy and effectiveness of medications for the treatment of accompanying somatization (KQ 3b) is low.

Table 25. Studies of adults with major depressive disorders and accompanying somatization

Study	Interventions	N	Results	Quality Rating
Kroenke et al., 2001[49]	Fluoxetine Paroxetine Sertraline	601	KQ 3b. Improvement in somatization scores was similar in all groups (*P* = NR)	Fair

NR, not reported.

Efficacy: KQ 3a: Depressive episode and KQ 3b: Accompanying symptoms. We identified no efficacy trials that addressed either of these questions in patients with depression and somatization.

Effectiveness: KQ 3a: Depressive episode in patients with somatization. We identified no trials addressing treatment of depression in subgroups of patients with somatization.

Effectiveness: KQ 3b: Somatization in depressed patients. One open-label study provided evidence for the comparative effectiveness of SSRIs for treatment of accompanying somatization in patients with depression.[49] This trial found no difference in effectiveness among paroxetine, fluoxetine, and sertraline on a somatization severity scale measure.

Somatization: Detailed Analysis

Head-to-head evidence. ***Fluoxetine vs. paroxetine vs. sertraline.*** One open-label, head-to head trial compared the effectiveness of low-dose fluoxetine, paroxetine, and sertraline for the treatment of depression in primary care over 9 months.[49] Somatization severity was measured using the Patient Health Questionnaire Somatization Severity scale (0-28 scale, where higher scores indicate worse severity). The report did not present analyses stratified by levels of somatization severity. The authors reported no statistically significant differences in somatization severity scores among treatment groups (-3.1 for fluoxetine, -3.2 for paroxetine, and -4.1 for sertraline, *P* = NR).

Placebo-controlled evidence. We identified no placebo-controlled trials addressing efficacy or effectiveness of treating patients with depression and somatization.

Key Question 4: Comparative harms and adherence for second-generation antidepressants

Overview

Most of the studies that examined the efficacy of one drug relative to another also determined differences in harms. Methods of adverse events assessment differed greatly. Few studies used objective scales such as the UKU-SES (Utvalg for Kliniske Undersogelser Side Effect Scale) or the adverse reaction terminology from the World Health Organization (WHO). Most studies combined patient-reported adverse events with a regular clinical examination by an investigator. Determining whether assessment methods were unbiased and adequate was often difficult. Rarely were adverse events prespecified and defined. Short study durations and small sample sizes also limited the validity of adverse events assessment in many trials.

Few RCTs were designed to assess adverse events as primary outcomes. Most published studies were post hoc analyses or retrospective reviews of databases. We included observational studies if the sample size was larger than 100 and the study duration was at least 3 months.

Detailed information on included studies can be found in Evidence Table 5 in Appendix D; information on systematic reviews and meta-analyses on this topic appears in Evidence Table 6. Most studies were rated fair quality; for three studies, a quality grade was not applicable because of the nature of the study design. Studies rated other than fair are noted in text.

Adverse Events and Discontinuation Rates: Key Points

We analyzed adverse events data of 72 head-to-head efficacy studies of 16,780 patients and 39 additional studies of both experimental and observational design. Of these, only five were designed primarily to detect differences in adverse events. The method of adverse events assessment in efficacy trials differed greatly. Few studies used objective scales. Determining whether assessment methods were unbiased and adequate was often difficult.

In efficacy trials, on average, 61 percent of patients experienced at least one adverse event during treatment. Nausea, headache, diarrhea, fatigue, dizziness, sweating, sexual dysfunction, tremor, dry mouth, and weight gain were the commonly reported adverse events. Overall, second-generation antidepressants led to similar adverse events.

However, results from multiple head-to-head RCTs document that the frequencies of *specific* adverse events can differ among drugs. These findings are generally consistent with results from observational studies. Specifically:

- Venlafaxine was associated with an approximately 10 percent (95% CI, 4-17 percent) higher incidence of nausea and vomiting than SSRIs as a class. In addition, pooled discontinuation rates because of adverse events in efficacy trials were statistically significantly higher for venlafaxine than for SSRIs (RR, 1.50; 95% CI, 1.21-1.84).

- In most studies, sertraline led to higher rates of diarrhea than comparator drugs (bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine).[30,42,44-46,51,53,76,77,90,92,108,110,111,150] Incidence was 8 percent (95% CI, 3-11

percent) higher than with comparator drugs. Whether this finding can be extrapolated to comparisons of sertraline with other second-generation antidepressants remains unclear.

- Mirtazapine led to higher weight gains than comparator drugs.[59-61,72,73,93,94] Mean weight gains relative to pretreatment weights ranged from 0.8 kg to 3.0 kg after 6 weeks to 8 weeks of treatment. Paroxetine had higher weight gains than fluoxetine and sertraline.[41,48]

- Trazodone was associated with an approximately 16 percent (3 percent less to 36 percent higher) higher incidence of somnolence than comparator drugs (bupropion, fluoxetine, mirtazapine, paroxetine, venlafaxine).[82,83,87,94,96,97] Whether this finding can be extrapolated to comparisons of trazodone with other second-generation antidepressants remains unclear.

Discontinuation syndromes (e.g., headache, dizziness, nausea) occurred in 0 percent to 86 percent of patients. Paroxetine and venlafaxine had the highest incidence of this problem, and fluoxetine had the lowest incidence.

Pooled estimates from efficacy trials suggest that these differences do not lead to any statistically significant differences in overall discontinuation rates among SSRIs as a class and other second-generation antidepressants.

Adverse Events and Discontinuation Rates: Detailed Analysis

Table 26 presents data on the design, interventions, results, and quality ratings of studies we included to examine issues relating to key adverse events and discontinuation. We focused on general tolerability and discontinuation (including nausea and vomiting and selected gastrointestinal problems), weight change, and discontinuation syndrome. We rated the strength of evidence on general adverse events as high or moderate (depending on the specific measure) and on discontinuation rates as high. Table 27 depicts, by specific drug, the mean incidence and 95 percent confidence interval for specific adverse events commonly reported in trials. Statistics are descriptive only. Comparisons across different drugs should be made with caution given differences in assessment and reporting of adverse events across trials.

General tolerability. Nausea, headache, diarrhea, fatigue, dizziness, sweating, tremor, dry mouth, and weight gain were commonly reported adverse events. In efficacy trials, on average, 61 percent of patients experienced at least one adverse event during the course of a given study.

Two large observational studies (three articles) examined the comparative rates of adverse events among SSRIs.[160-162] Overall, no substantial differences among examined drugs were apparent. However, not all currently approved SSRIs were investigated in these studies.

A British study pooled data from a cross-sectional study of a prescription-event monitoring study of general practitioners 6 months to 1 year after they had issued prescriptions.[160,161] Included drugs were fluoxetine, fluvoxamine, paroxetine, sertraline, venlafaxine, and nefazodone. The final cohort exceeded 10,000 patients for each drug. Demographics and

Table 26. Studies assessing adverse events and discontinuation rates

Study	Design Interventions	N	Results	Quality Rating
General Tolerability and Discontinuation				
Brambilla et al., 2005[163]	Systematic review, Fluoxetine vs. SSRIs	NR	No difference in discontinuation rates because of adverse events	Good
Greist et al., 2004[164]	Pooled analysis Duloxetine vs. paroxetine vs. fluoxetine	2,345	No differences in nausea between duloxetine and paroxetine or duloxetine and fluoxetine	NA
Haffmans et al., 1996[29]	RCT Fluvoxamine vs. paroxetine	217	Significantly more diarrhea and nausea with fluvoxamine	Fair
Mackay et al., 1997,[160] 1999[161,165]	Prescription event monitoring Fluoxetine, fluvoxamine, nefazodone, paroxetine, venlafaxine	> 60,000	Venlafaxine had highest rate of nausea and vomiting; fluvoxamine had the most overall adverse events	NA
Meijer et al., 2002[162]	Observational study Sertraline vs. SSRIs	1,251	Significantly more diarrhea with sertraline	Fair
Rapaport et al., 1996[33]	RCT Fluoxetine vs. fluvoxamine	100	Significantly more nausea with fluoxetine	Fair
Changes in Weight				
Benkert et al., 2000[72]	RCT Paroxetine vs. mirtazapine	275	Higher weight gain with mirtazapine	Fair
Croft et al., 2002[166]	RCT Bupropion vs. placebo	423	Small weight loss with bupropion over 44 weeks	Fair
Fava et al., 2000;[48] Fava et al., 2002[41]	RCT Fluoxetine vs. paroxetine vs. sertraline	284	Highest weight gain with paroxetine	Fair
Goldstein et al.,1997[167]	RCT Fluoxetine vs. placebo	671	Higher weight loss with fluoxetine in older patients	Fair
Guelfi et al., 2001[93]	RCT Venlafaxine vs. mirtazapine	157	Higher weight increase with mirtazapine	Fair
Halikas et al., 1995[94]	RCT Trazodone vs. Mirtazapine	150	More weight gain with mirtazapine	Fair
Harto et al., 1988[168]	RCT Fluoxetine vs. placebo	35	Higher weight loss with fluoxetine	Fair
Hong et al., 2003[59]	RCT Fluoxetine vs. mirtazapine	133	Higher weight gain with mirtazapine	Fair
Michelson et al., 1999[145] Reimherr et al., 1998[122]	RCT Fluoxetine vs. placebo	395	Fluoxetine and placebo showed a weight gain	Fair
Schatzberg et al., 2002[73]	RCT Paroxetine vs. mirtazapine	255	Higher weight gain with mirtazapine	Fair

NA, not applicable; NR, not reported; RCT, randomized controlled trial.

Table 26. Studies assessing adverse events and discontinuation rates (continued)

Study	Design Interventions	N	Results	Quality Rating
Versiani et al., 2005 [60]	RCT Fluoxetine vs. mirtazapine	297	Higher weight gain with mirtazapine	Fair
Wheatley et al., 1998[61]	RCT Fluoxetine vs. mirtazapine	133	Significantly higher weight gain with mirtazapine	Fair
Discontinuation Syndrome				
CSM Expert Working Group, 2004[169]	Systematic review and meta-analysis Second-generation antidepressants	NR	No differences in risk among second-generation antidepressants	Good
Judge et al., 2002[170]	Open-label trial Fluoxetine and paroxetine	150	Significantly fewer symptoms in the fluoxetine group than the paroxetine group	Fair
Perahia et al., 2005[171]	Pooled analysis Duloxetine vs. placebo	3,624	Significantly higher rate of discontinuation syndrome with duloxetine than with placebo (44% vs. 23%)	Fair
Zajecka et al., 1998[172]	RCT Fluoxetine vs. placebo	395	Dizziness significantly less frequent in fluoxetine patients at 4 and 6 weeks	Fair

Table 27. Mean incidence of specific adverse events across comparative trials

Drug	Diarrhea	Dizziness	Headache	Insomnia	Nausea	Somnolence
Mean Percentage* (95% confidence interval)						
Bupropion	10.2% (3.1%-17.2%)	11.6% (2.2%-21.1%)	28.6% (23.2%-34.1%)	15.7% (10.9%-20.6%)	14.5% (8.9%-20%)	6.7% (0%-14.2%)
Citalopram	7.5% (4%-11%)	9.1% (3.7%-14.4%)	14.3% (7.8%-20.7%)	6.9% (1.4%-12.5%)	14.3% (9.6%-19.1%)	12.6% (5.4%-19.9%)
Duloxetine	16.1% (9.5%-22.8%)	41.5% (-8.1%-91%)	15.8% (3.9%-27.7%)	16.6% (14.1%-19.1%)	42.6% (7.2%-78%)	36.8% (8.4%-65.2%)
Escitalopram	7.6% (0%-16%)	1.3% (0%-14.3%)	7.4% (3.3%-11.5%)	6.9% (1.3%-10.8%)	11.5% (7.2%-15.7%)	4.2% (0%-12.2%)
Fluoxetine	10.4% (7.5%-13.3%)	7.6% (6.2%-9%)	21.3% (16.3%-26.3%)	13.8% (11.4%-16.2%)	18.4% (15.9%-20.9%)	7.8% (5.3%-10.3%)
Fluvoxamine	19.2% (0%-53.5%)	18.3% (0%-62.4%	20.1% (3.3%-36.8%)	24.2% (0.3%-48%)	26% (14.4%-37.6%)	8.8% (0%-32.2%)
Mirtazapine	3.7% (0%-8.1%)	8.4% (4.6%-12.1%)	12.1% (10%-14.3%)	8% (1.8%-14.3%)	6.3% (3.8%-8.7%)	18.7% (10.3%-27.1%)
Nefazadone	12% (7.3%-16.8%)	21.3% (15.6%-27%)	32.4% (21.6%-43.2%)	13.3% (7%-19.5%)	21.6% (12.2%-30.9%)	25.3% (11.4%-39.1%)
Paroxetine	15% (11.1%-18.9%)	0.8% (0%-2.9%)	3.2% (0%-8.1%)	12.7% (9.9%-15.4%)	21.4% (17.1%-25.7%)	18.2% (13.7%-22.7%)
Sertraline	11.3% (7.6%-15%)	8.5% (5.9%-11.2%)	19.8% (14.9%-24.7%)	9.8% (6.1%-13.6%)	17.3% (13.7%-20.8%)	13.3% (9.8%-16.8%)
Trazodone	4.3% (0%-13.8%)	24.1% (11.8%-36.5%)	22.1% (11.7%-32.5%)	4.8% (1.8%-7.8%)	14.4% (4.6%-24.1%	42.4% (19.5%-65.2%)
Venlafaxine	6.4% (2.9%-10%)	14.3% (8.9%-19.7%)	19.3% (13.9%-24.7%)	17.8% (12.2%-23.2%)	29.3% (24.8%-33.8%)	14.5% (9.5%-19.4%)

*Weighted mean incidence calculated from RCTs. Method and extent of adverse event assessment varied among studies. Comparisons across drugs must be made cautiously.

indications were similar among study groups. Overall, the mean incidence per 1,000 patient-months for SSRIs was highest for fluvoxamine (fluvoxamine, 17.6; fluoxetine, 7.0; paroxetine, 7.6; sertraline, 6.2). Physicians, not patients, reported adverse events; the nonresponse rate was 40 percent. Therefore, measurement bias, selection bias, and potential confounding may compromise these results.

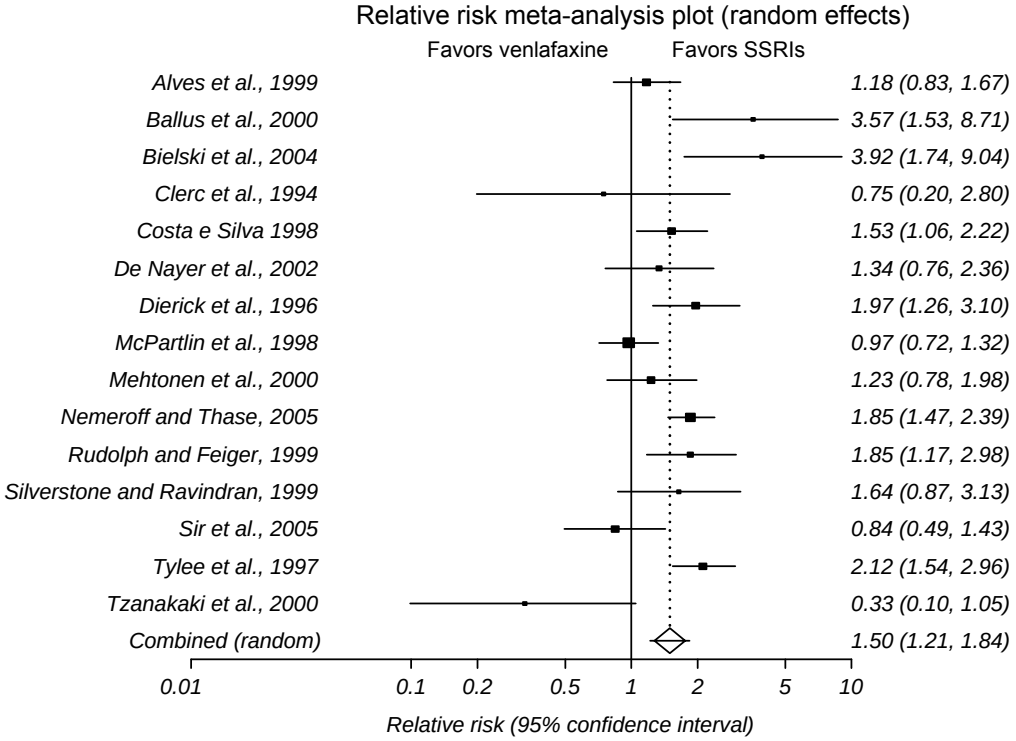
A Dutch prospective observational study followed 1,251 patients for up to 12 months to assess adverse events of sertraline (n = 659) and other SSRIs (fluoxetine, fluvoxamine, paroxetine).[162] No exclusion criteria were applied. Psychiatrists recorded adverse events at each patient visit; the investigators used the WHO adverse reaction terminology for outcome assessment. Significantly more sertraline patients than patients on other drugs had a diagnosis of depressive disorder at baseline ($P < 0.001$). Overall, 74.1 percent of patients reported at least one adverse event.

Nausea and vomiting. In efficacy trials, venlafaxine (an SNRI) had a consistently higher rate of nausea and vomiting than SSRIs. In six studies, the difference reached statistical significance.[56,57,65,67,70,74] The rate of patients reporting nausea or vomiting ranged from 6 percent to 48 percent.

These findings are consistent with a British prescription-event monitoring study described earlier.[160,161] Nausea and vomiting were the two most frequent clinical reasons for withdrawal in the first month of treatment for all drugs. Venlafaxine had the highest rate of nausea and vomiting per 1,000 patient-months.

Using data from efficacy trials, we compared the pooled relative risk of nausea and vomiting for venlafaxine with that for SSRIs as a class (Figure 14). The RR was 1.50 (95% CI, 1.21-1.84). The corresponding number needed to harm (NNH) was 9 (95% CI, 6-23).

Figure 14. Relative risk of nausea and vomiting of venlafaxine compared with SSRIs



A pooled analysis of published and unpublished trials of duloxetine did not find significant differences in nausea between duloxetine (40-120 mg/day) and paroxetine (20 mg/day) or between duloxetine (120 mg/day) and fluoxetine (20 mg/day).[164]

Changes in weight. Consistently, studies comparing mirtazapine with other second-generation antidepressants reported higher weight gains for mirtazapine than for the comparator groups.[59-61,72,73,93,94] In two RCTs, these differences reached statistical significance.[72,73] Mean weight gains ranged from 0.8 kg to 3.0 kg after 6 to 8 weeks of treatment. Standard deviations of these changes, however, were large, suggesting that some patients had substantially higher weight increases.

Three placebo-controlled RCTs specifically assessed weight changes with fluoxetine treatment.[122,145,167,168] Findings were mixed. Two studies, one conducted in 671 patients older than 60 years,[167] recorded a statistically significant weight loss for fluoxetine compared with placebo.[167,168] The third study reported a weight gain.[122,145]

A 32-week acute- and continuation-phase trial assessed differences in weight changes among patients treated with fluoxetine, paroxetine, and sertraline.[41,48] Paroxetine patients showed a significantly greater mean weight change (+3.6 percent) than those taking fluoxetine (-0.2 percent; $P = 0.015$) and sertraline (+1.0 percent; $P < 0.001$). With respect to weight gain of more than 7 percent, significantly more patients in the paroxetine group (25.5 percent) than in the fluoxetine group (6.8 percent; $P = 0.016$) and the sertraline group (4.2 percent; $P = 0.003$) had weight gains of this magnitude.

A double-blinded, placebo-controlled, 52-week acute- and continuation-phase trial assessed weight changes during bupropion treatment.[166] Patients receiving bupropion showed a modest

but nevertheless significant decrease in body weight from baseline (-1.15 kg; $P < 0.001$). The magnitude of weight change was closely related to the patient's body mass index (BMI). Patients with a higher BMI experienced greater weight loss.

Gastrointestinal adverse events. Two RCTs were designed primarily to detect differences in harms between fluvoxamine and citalopram[29] and fluvoxamine and fluoxetine.[33] A Dutch multicenter trial assessed gastrointestinal side effects from citalopram (20-40 mg/day) and fluvoxamine (100-200 mg/day).[29] A total of 217 patients were enrolled for 6 weeks. Overall, 57 percent of patients reported adverse events. Significantly more patients in the fluvoxamine group than in the citalopram group had diarrhea (+13 percent; $P = 0.026$) or nausea (+16 percent; $P = 0.017$). However, the authors did not provide a baseline comparison of gastrointestinal illnesses between groups, so differences at baseline could bias results.

Another trial assessed differences in adverse events between fluvoxamine (100-150 mg/day) and fluoxetine (20-80 mg/day) in 100 patients over 7 weeks.[33] No significant difference could be detected, except that patients on fluoxetine suffered nausea significantly more often than those on fluvoxamine (42.5 percent vs. NR; $P = 0.03$).

In a Dutch prospective observational study (N = 1,251), diarrhea occurred more frequently in the sertraline group than in patients on fluoxetine, fluvoxamine, and paroxetine ($P < 0.05$).[162] This finding is consistent with results from head-to-head efficacy studies. In most studies, sertraline led to higher rates of diarrhea than did comparator drugs (bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, and venlafaxine).[30,42,44-46,51,53,76,77,90,92,108,110,111,150] Incidence was 8 percent (95% CI, 3-11 percent) higher than with comparator drugs. The NNH is 13 (95% CI, 9-29). Whether this finding can be extrapolated to comparisons of sertraline with other second-generation antidepressants remains unclear.

Discontinuation syndrome. Withdrawal syndromes (e.g., headache, dizziness, lightheadedness, nausea, anxiety) commonly occur following the abrupt discontinuation of second-generation antidepressants. A good systematic review conducted by an Expert Working Group of the UK Committee on Safety in Medicines (CSM) assessed the frequency of discontinuation syndromes in second-generation antidepressants.[169] Based on observational studies, spontaneous reporting data, and clinical trials data, discontinuation syndromes occurred in 0 percent to 86 percent of patients. Because of study durations, dosages, and different assessment methods, incidence rates could not be compared directly. Nevertheless, discontinuation syndromes occurred most commonly with paroxetine and venlafaxine and least commonly with fluoxetine.

Three studies not included in the UK systematic review provide consistent results.[170-172] One head-to-head trial compared fluoxetine with paroxetine.[170] Treatment interruption led to significantly fewer symptoms in the fluoxetine group than the paroxetine group ($P = 0.001$) using the Discontinuation-Emergent Signs and Symptoms checklist (DESS). A placebo-controlled trial of fluoxetine did not find any differences in discontinuation syndromes between fluoxetine and placebo.[172] A pooled analysis of six trials investigated the effects of abrupt discontinuation of duloxetine and placebo.[171] Significantly more patients receiving duloxetine than receiving placebo reported discontinuation syndromes (44.3 percent vs. 22.9 percent; $P < 0.05$).

Discontinuation rates. In efficacy trials, discontinuation rates because of adverse events were not substantially different. Using data from efficacy studies, we conducted meta-analyses to assess differences in the overall loss to followup, discontinuation rates because of adverse events, and discontinuation rates because of lack of efficacy of SSRIs as a class compared with other second-generation antidepressants (bupropion, duloxetine, mirtazapine, nefazodone, trazodone, and venlafaxine) in adult patients with MDD.

Table 28 summarizes average discontinuation rates. Figures 15 through 17 depict relative risks of discontinuation rates comparing these agents with SSRIs as a class. Three sets of individual meta-analyses for overall discontinuation and for discontinuation from adverse events and lack of efficacy are presented in Appendix H. Available data for duloxetine and trazodone were insufficient to determine discontinuation rates that might be attributed to lack of efficacy.

Overall discontinuation rates did not differ significantly between SSRIs and bupropion, duloxetine, mirtazapine, nefazodone, trazodone, or venlafaxine (Figure 15). A published meta-analysis of 15 RCTs did not find any statistically significant differences in discontinuation rates because of adverse events between fluoxetine and other SSRIs as a class.[163] The only statistically significant difference in our pooled estimates of the relative risks of discontinuation because of

Table 28. Average rates of overall discontinuation, discontinuation because of adverse events, and discontinuation because of lack of efficacy

	Overall Loss to Followup (%)	Discontinuation Because of Adverse Events (%)	Discontinuation Because of Lack of Efficacy (%)
SSRIs	20.8	8.1	4.4
Bupropion	14.1	6.7	3.1
Duloxetine	17.2	5.5	NR
Mirtazapine	21.6	9.5	3.4
Nefazodone	23.6	15.0	2.0
Trazodone	20.7	7.0	NR
Venlafaxine	24.8	11.5	3.5

Figure 15. Relative risks of overall discontinuation

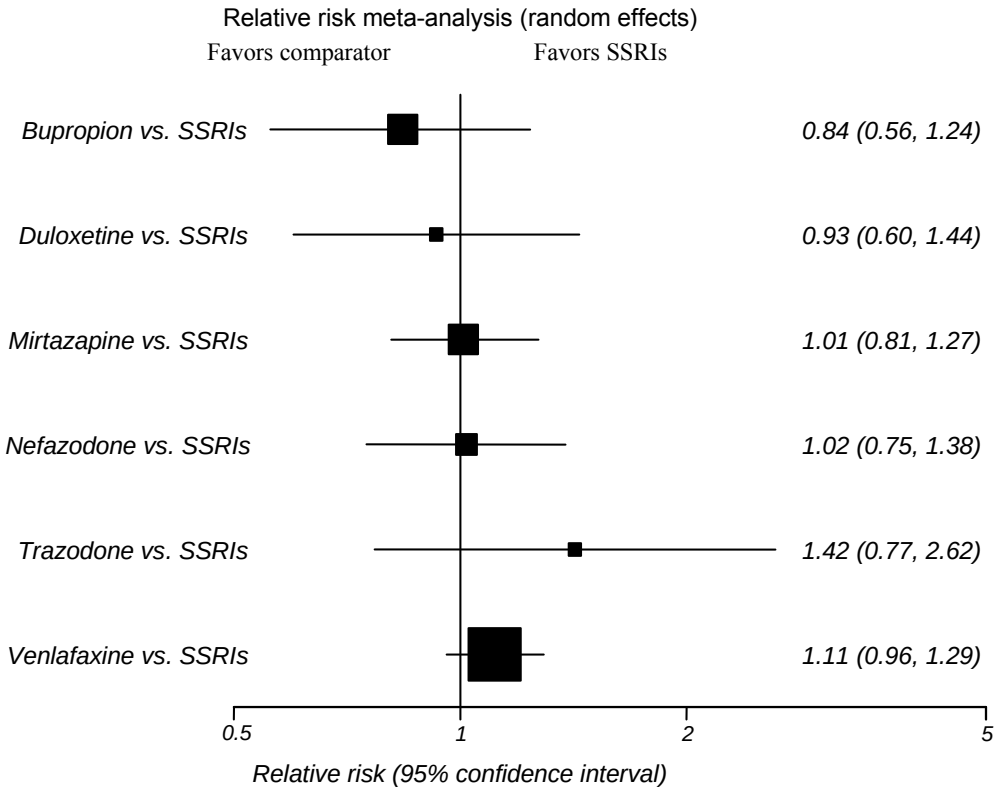


Figure 16. Relative risk of discontinuation because of adverse events

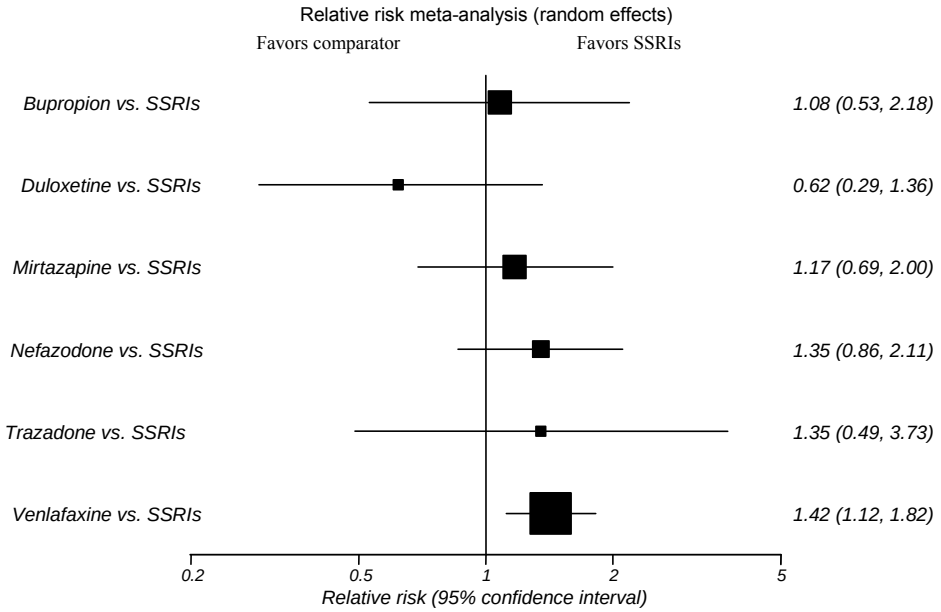
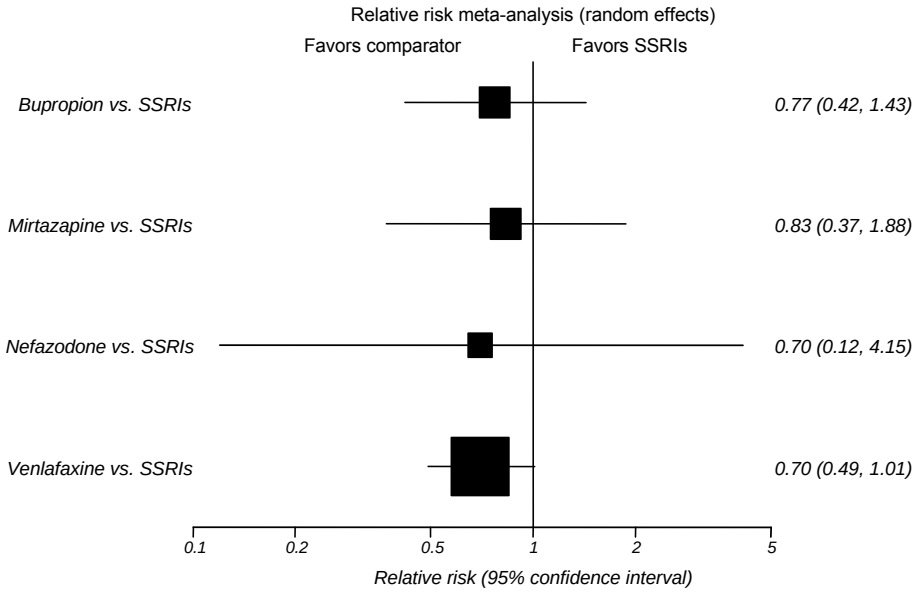


Figure 17. Relative risk of discontinuation because of lack of efficacy



adverse events was a higher rate for patients on venlafaxine than for patients on SSRIs (RR, 1.42; 95% CI, 1.12-1.82) (Figure 16). Overall, this finding was balanced by lower discontinuation rates because of lack of efficacy for venlafaxine (RR, 0.70; 95% CI, 0.49-1.01) (Figure 17).

Nefazodone and trazodone had rates of discontinuation because of adverse events similar to those of venlafaxine. However, differences with SSRIs did not reach statistical significance because of smaller sample sizes.

Severe Adverse Events: Key Points

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

Based on five RCTs (N = 1,489), bupropion led to a significantly lower rate of sexual adverse events than fluoxetine and sertraline.[79,80,88,89,102] The NNT to experience one additional person with high overall satisfaction of sexual functioning is 6 (95% CI, 4-9).

Compared with other second-generation antidepressants (fluoxetine, fluvoxamine, nefazodone, and sertraline), paroxetine frequently led to higher rates of sexual dysfunction (16 percent vs. 6 percent).[37,41,50,86] Underreporting of absolute rates of sexual dysfunction, however, is likely in these studies. The NNH is 6 (95% CI, 4-9).

Table 29 summarizes studies included for the assessment of severe adverse events: suicidality, sexual dysfunction, cardiovascular events, seizures and other events.[173]

Table 29. Studies assessing severe adverse events

Study	Design Interventions	N	Results	Quality Rating
Suicidality (Suicidal thoughts and behavior)				
CSM Expert Working Group, 2004[169]	Systematic review and meta-analysis Second-generation antidepressants	NR	No differences in risk among second-generation antidepressants	Good
Fergusson et al., 2005[174]	Meta-analysis SSRIs vs. placebo	87,650	Higher risk of suicide attempts for SSRI-treated patients	Good
Khan et al., 2003[175]	Retrospective cohort study Bupropion, citalopram, fluoxetine, mirtazapine, nefazodone, paroxetine, venlafaxine	48,277	No difference in the rate of suicides	Fair
Gunnell et al., 2005[176]	Meta-analysis Citalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, all vs. placebo	40,000	Increased risk of nonfatal suicide attempts compared with placebo; no difference in risk among drugs	Good
Martinez et al., 2005[177]	Case-control study Citalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, TCAs	146,095	No difference in risk of suicide or nonfatal suicide attempts between SSRIs and TCAs or among individual SSRIs	Good
Didham et al., 2005[178]	Retrospective cohort study Citalopram, fluoxetine, paroxetine	57,000	Significant association between nonfatal suicide attempts and SSRIs; no difference in risk among drugs	Fair
Pedersen, 2005[173]	Retrospective cohort study Escitalopram vs. placebo	4,091	Higher rate of nonfatal suicide attempts for escitalopram than for placebo	Fair
Jick et al., 1992[179]	Database review Fluoxetine, first-generation antidepressants	8,730	No difference in suicides between fluoxetine and first-generation antidepressants	NA
Jick et al., 2004[180]	Case-control study Fluoxetine, paroxetine	159,810	No difference in risk among drugs	Fair
Jick et al., 1995[181]	Retrospective cohort study and nested case-control study Fluoxetine, trazodone, first-generation antidepressants	172,598	Significantly higher risk of suicide for fluoxetine and mianserin than for dothiepin	Fair
Aursnes et al., 2005[182]	Meta-analysis of unpublished data Paroxetine	1,466	Higher rate of suicides for paroxetine than for placebo	Fair
Lopez-Ibor, 1993[183]	Database review Paroxetine, first-generation antidepressants	4,686	No difference in suicidality	NA

CSM, Committee on Safety in Medicines; NA, not applicable; NR, not reported; RCT, randomized controlled trials; TCAs, tricyclic antidepressants.

Table 29. Studies assessing severe adverse events (continued)

Study	Design, Interventions	N	Results	Quality Rating
Sexual Dysfunction				
Clayton et al., 2002[184]	Cross-sectional survey Bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine	6,297	Highest risk for paroxetine, lowest risk for bupropion	Fair
Coleman et al., 2001[79]	RCT Bupropion SR vs. fluoxetine	456	Significantly more sexual adverse events with fluoxetine	Fair
Feighner et al., 1991[80]	RCT Bupropion vs. fluoxetine	61	Higher rate of sexual dysfunction for fluoxetine	Fair
Coleman et al., 1999[88]	RCT Bupropion SR vs. sertraline	364	Significantly more sexual adverse events with sertraline	Fair
Croft et al., 1999[89]	RCT Bupropion SR vs. sertraline	360	No differences	Fair
Segraves et al., 2000[102]	RCT Bupropion vs. sertraline	248	Significantly more sexual adverse events with sertraline	Fair
Nieuwstraten and Dolovich, 2001[191]	Meta-analysis Bupropion vs. SSRIs	1,332	Significantly higher rate of sexual satisfaction in bupropion group	Good
Montejo et al., 2001[190]	Prospective cohort study Citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine	1,022	Highest incidence of sexual dysfunction for citalopram, paroxetine, and venlafaxine; lowest for mirtazapine and nefazodone	Fair
Landen et al., 2005[189]	Cross-sectional study Citalopram vs. paroxetine	119	No differences	Fair
Ekselius et al., 2001[186]	RCT Citalopram vs. sertraline	308	No differences	Fair
Delgado et al., 2005[185]	Pooled analysis Duloxetine vs. paroxetine vs. placebo	1,466	Higher rate of sexual dysfunction for paroxetine	Fair
Philip et al., 2000[192]	Prospective cohort study Fluoxetine, fluvoxamine, paroxetine, sertraline, moclobemide	268	No difference among SSRIs	Fair
Fava et al., 1998[37]	Pooled Analysis Fluoxetine vs. paroxetine	128	Significantly more sexual adverse events with paroxetine	Fair
Kennedy et al., 2000[188]	Prospective cohort study Paroxetine, sertraline, venlafaxine	174	No difference	Fair
Kavoussi et al., 1997[90,108]	RCT Sertraline vs. bupropion	248	Higher rate of sexual adverse events with sertraline	Fair
Nemeroff et al., 1995[51]	RCT Sertraline vs. fluvoxamine	95	Higher rate of sexual adverse events with sertraline	Fair
Behnke et al., 2003[76]	RCT Sertraline vs. mirtazapine	346	Significantly more sexual adverse events with sertraline	Fair
Feiger et al., 1996[92]	RCT Sertralinenefazodone	160	Sertraline had significant adverse effects on sexual function; nefazodone had none	Fair
Aberg-Wistedt et al., 2000[53]	RCT Sertralineparoxetine	353	Significantly more libido decreases in patients taking sertraline	Fair
Ferguson et al., 2001[187]	RCT Sertraline vs. trazodone	150	Higher reemergence rate of sexual dysfunction for sertraline	Fair

Table 29. Studies assessing severe adverse events (continued)

Study	Design, Interventions	N	Results	Quality Rating
Seizures				
Dunner et al., 1998[193]	Uncontrolled, open-label trial Bupropion	3,100	Rate of seizures for bupropion within reported range of other antidepressants	Fair
Johnston et al., 1991[194]	Uncontrolled, open-label trial Bupropion	3,341	Rate of seizures for bupropion within range of other antidepressants	Fair
Whyte et al., 2003[195]	Prospective observational study SSRIs, TCAs, venlafaxine	538	Seizures more common in venlafaxine overdose than in SSRI or TCA overdose	Good
Cardiovascular Events				
Thase et al., 1998[196]	Pooled analysis Venlafaxine	3,744	Increase in diastolic blood pressure for venlafaxine	Fair
Thase et al., 2005[197]	Post hoc data analysis Fluoxetine, paroxetine, duloxetine	1,873	Greater change in heart rate for duloxetine than for fluoxetine and paroxetine	NA
Other Adverse Events				
Buckley et al., 2002[198]	Database analysis Citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, venlafaxine	47,329	Highest rate of fatal toxicity for venlafaxine	NA
Coogan et al., 2005[199]	Case-control SSRIs	4,996	No association between breast cancer and SSRIs	Fair
Kirby et al., 2002[200]	Retrospective cohort study SSRIs, venlafaxine	199	Increased rate of hyponatremia in patients on SSRIs and venlafaxine	Fair
Thapa et al., 1998[201]	Retrospective cohort study Fluoxetine, paroxetine, sertraline, trazodone	2,428	No difference in the risk of falls	Fair

Severe Adverse Events: Detailed Analysis

Suicidality. Eleven studies (12 articles) assessed the risk of suicidality (suicidal thinking or behavior) in patients treated with second-generation antidepressants.[169,174-183] Data on the comparative risk of suicidality among second-generation antidepressants are sparse. Results from existing studies do not indicate that any particular drug of interest has an excess risk compared with that of other second-generation antidepressants.[175-178,180] However, these findings are based primarily on retrospective cohort studies,[175,177,178,180] and confounding by indication (i.e., patients who are at higher risk for suicide may be prescribed some medications rather than others) may lead to erroneous conclusions.

The largest attempt to determine whether second-generation antidepressants increase the risk of suicidality was conducted in 2004 by the CSM working group.[169] The CSM experts investigated ongoing safety concerns about suicidal behavior with some second-generation antidepressants (citalopram, escitalopram, fluoxetine, fluvoxamine, mirtazapine, paroxetine, sertraline, venlafaxine) in patients with MDD. They used data from 477 published and unpublished RCTs on more than 40,000 individuals as well as spontaneous reporting data. However, these data were limited to studies funded by the pharmaceutical industry.

A meta-analysis limited the CSM data to placebo-controlled trials of SSRIs in ~40,000 adults. Results did not yield any evidence that SSRIs either increase or protect against the risk of suicide (OR, 0.85; 95% CI, 0.20-3.40).[176] The risk of suicide-related events was similar between second-generation antidepressants and active comparators, although some evidence of an increased risk of nonfatal suicide attempts was detected (OR, 1.57; 95% CI, 0.99-2.55).

Another meta-analysis of published data on more than 87,000 patients in SSRI trials for various conditions reported a significantly higher risk of suicide attempts for SSRI patients than for placebo-treated patients (OR, 2.25; 95% CI, 1.14-4.55).[174] Furthermore, an increase in the odds ratio of suicide attempts was observed for SSRIs compared with interventions other than tricyclic antidepressants (TCAs) (OR, 1.94; 95% CI, 1.06-3.57). No significant difference existed in the pooled analysis of SSRIs compared with TCAs (OR, 0.88; 95% CI, 0.54-1.42). The overall rate of suicide attempts was 3.9 (95% CI, 3.3-4.6) per 1,000 patients treated with SSRIs, for an incidence of 18.2 suicide attempts per 1,000 patient years.

In addition, the CSM group commissioned an observational study (i.e., a nested case-control study) using the General Practice Research Database (GPRD) to investigate the association between antidepressants and suicide attempts. This study used data on more than 146,000 patients with a first prescription of an antidepressant for depression.[177] It did not find any evidence that the risk of either suicide (OR, 0.57; 95% CI, 0.26-1.25) or nonfatal suicide attempts (OR, 0.99; 95% CI, 0.86-1.14) was greater in patients on second-generation antidepressants than in patients on TCAs.

Findings of other large observational studies and meta-analyses are similar.[173,175,178-183] Most detected an increase in nonfatal suicide attempts but no significant difference in suicides. In general, no significant differences in risks regarding suicidality could be detected between second-generation antidepressants and TCAs.

An internal report of all published and unpublished studies on paroxetine conducted by GlaxoSmithKline is consistent with findings from studies described above.[202]

Sexual dysfunction. Multiple studies assessed the comparative risk of sexual dysfunction among second-generation antidepressants.[79,88,89,102,190] The largest study was a Spanish open-label, prospective observational study using the Psychotropic-Related Sexual Dysfunction Questionnaire (PRSexDQ) in 1,022 outpatients treated with various antidepressants.[190] All patients had normal sexual functioning at study onset. Overall, 59 percent of patients experienced some type of sexual dysfunction. Among second-generation antidepressants, citalopram, paroxetine, and venlafaxine had the highest incidence of sexual dysfunction (73 percent, 71 percent, and 67 percent, respectively); mirtazapine and nefazodone had the lowest (24 percent and 8 percent, respectively). This study did not include data on bupropion, escitalopram, and trazodone.

A cross-sectional survey of patients on second-generation antidepressants presented similar results.[184] Paroxetine had the highest rate of sexual dysfunction; nefazodone and bupropion had the lowest.

Sexual dysfunction was also a commonly reported adverse event for SSRIs and SNRIs in efficacy trials. Most of these studies did not report the use of targeted questions for sexual dysfunction. Therefore, patient-reported numbers might not reflect the true incidence. Patients receiving paroxetine and sertraline frequently reported significantly higher rates of sexual dysfunction[37,51,53,76,90,92] than did patients in the active control groups. In one trial, significantly more patients on sertraline than on bupropion SR withdrew because of sexual dysfunction (13.5

percent vs. 3.3 percent; $P = 0.004$).[90] A pooled analysis of four efficacy trials comparing paroxetine and duloxetine reported significantly higher rates of sexual dysfunction for patients on paroxetine.[185]

Six RCTs assessed the comparative risk of sexual dysfunction between two or more second-generation antidepressants as primary outcome measures.[79,88,89,102,186,187]

Citalopram vs. sertraline. A subgroup analysis of a Swedish RCT examined the incidence of sexual dysfunction from citalopram (20-60 mg/day) and from sertraline (50-150 mg/day) in 308 study completers with MDD.[186] Outcome assessment was conducted at baseline and at week 24. Citalopram and sertraline did not differ significantly in the magnitude and frequency of sexual dysfunction. Only one patient was lost to followup attributable to sexual dysfunction in this study.

Bupropion vs. SSRIs. A good meta-analysis including data on 1,332 patients with MDD compared sexual adverse events of bupropion and three SSRIs (fluoxetine, paroxetine, sertraline) as a class.[191] We do not describe studies included in this meta-analysis individually.[80,88,89,102] The rate of sexual satisfaction was significantly higher in patients receiving bupropion than in those receiving SSRIs (RR, 1.28; 95% CI, 1.16-1.41).

An 8-week RCT (not in the meta-analysis cited above) compared efficacy and sexual dysfunction of bupropion SR (150-400 mg/day), fluoxetine (20-60 mg/day), and placebo in 456 outpatients with MDD.[79] Findings are consistent with those from the meta-analysis. Throughout the study, patients on bupropion SR experienced significantly less sexual dysfunction than those on fluoxetine. Moreover, beginning at week 1 until endpoint, significantly fewer patients on bupropion than on fluoxetine were dissatisfied with their overall sexual function ($P < 0.05$). The NNT to gain one more patient with high satisfaction with sexual functioning is 6 (95% CI, 4-9).

Sertraline vs. trazodone. In one RCT, the emergence of sexual adverse events in patients who experienced sexual dysfunction with sertraline treatment was significantly greater for those receiving sertraline than for those receiving trazodone.[187]

Seizures. Evidence from controlled trials and observational studies is insufficient to conclude for or against an increased risk of seizures in patients taking any of the reviewed drugs, including bupropion. Two open-label trials[193,194] examined the rate of seizures during bupropion treatment. Both trials reported that the rate of seizures was within the range of other marketed antidepressants, but we rate the strength of this uncontrolled, open-label evidence as low. A recent review of medical charts on 538 patients with deliberate self-poisoning with antidepressants reported that seizures were more common in patients with venlafaxine overdose than in patients with TCA or SSRI overdose.[195]

Cardiovascular adverse events. A post hoc analysis examined pooled data from 3,744 patients participating in venlafaxine trials.[196] At 6 weeks, 11.5 percent of venlafaxine patients had a supine diastolic blood pressure (DBP) greater than 90 mm Hg (placebo, 5.7 percent; $P < 0.001$). During continuation-phase treatment (up to 12 months), significantly more venlafaxine subjects than placebo subjects with normal supine DBPs developed elevated readings ($P = 0.05$).

A post hoc analysis of six RCTs (published and unpublished) comparing duloxetine with fluoxetine and paroxetine did not find any statistically significant differences in supine systolic or diastolic blood pressure.[197] Patients receiving duloxetine had a greater mean change in heart rate than those on either fluoxetine (+2.8 beats/min vs. -1.0 beat/min) or paroxetine (+1.0 beats/min vs. -1.4 beats/min).

Efficacy trials infrequently assessed cardiovascular outcomes. Two RCTs, one comparing venlafaxine XR with sertraline[78] and one comparing venlafaxine with fluoxetine,[66] detected statistically significant increases in supine DBP[78] and supine pulse rate[66] for venlafaxine relative to fluoxetine.

Other adverse events. A database analysis in the United Kingdom on fatal toxicity of second-generation antidepressants found venlafaxine to have the highest fatal toxicity rate (13.2 per 1,000,000 prescriptions) among second-generation antidepressants.[198] A case-control study did not find an association between SSRIs and breast cancer.[199] A retrospective review of the charts of 2,428 nursing home residents did not detect differences in the risk of falls among fluoxetine, paroxetine, and sertraline.[201]

Hyponatremia. A retrospective cohort study reported that hyponatremia in elderly inpatients (mean age 74 years) was significantly more common in patients treated with SSRIs or venlafaxine than in controls not on these drugs (OR, 3.5; 95% CI, 1.4-8.9).[200] Otherwise, evidence from controlled trials and observational studies is insufficient to conclude for or against an increased risk of hyponatremia in patients treated with SSRIs.

Our methods for this comparative effectiveness review did not permit inclusion of case reports and case series. The published literature includes numerous case reports of hyponatremia and inappropriate secretion of an antidiuretic hormone as rare side effects.[203] Even if this evidence is considered weak, such findings might be important in the absence of studies with the methodological strength to account for rare adverse events.

Hepatotoxicity. Evidence from controlled trials and observational studies is insufficient to conclude for or against an increased risk of liver toxicity during nefazodone treatment. Nevertheless, numerous case reports not included in this report contain low quality but potentially important evidence citing an increased risk of liver toxicity during nefazodone treatment.[204]

Serotonin syndrome. Serotonin hyperstimulation syndrome is characterized by symptoms that include mental status changes, agitation, myoclonus, hyperreflexia, sweating, shivering, tremor, diarrhea, lack of coordination, and fever; it can lead to death.[165] Evidence from controlled trials and observational studies is insufficient to draw conclusions about differences in risk among second-generation antidepressants. The published literature has numerous case reports of serotonin syndrome.[205]

A postmarketing survey identified cases of the serotonin syndrome in British general practice among patients who received nefazadone.[165] In a cohort of 11,834 patients, 19 cases met criteria for the syndrome (incidence = 0.4 cases per 1,000 patient-months of treatment with nefazodone). Similar rates of the syndrome were reported for fluoxetine, sertraline, paroxetine, and venlafaxine.

Adherence: Key Points

Few efficacy studies reported rates of adherence. Lack of adherence, however, was often used as a reason to exclude patients from the study. Efficacy trials do not indicate any differences in adherence among second-generation antidepressants. However, the quality of reporting and assessment of adherence was limited. Findings from highly controlled efficacy studies may have limited generalizability to "real-world" practice especially because of the overall short duration of these trials. The evidence is insufficient to conclude on adherence in effectiveness studies. A review of a large, managed care database suggested that extended-release formulations might have greater adherence than immediate-release medications.[206] Strength of evidence is moderate for efficacy studies and low for effectiveness studies.

Adherence: Detailed Analysis

The published literature in this area frequently uses the terms "compliance" and "adherence" interchangeably. Compliance has traditionally been used to describe a patient's ability to take medications as prescribed. Some authors argue, however, that adherence better represents the more complex relationship among patients, providers, and medications; it is meant to reflect the fact that following a medication regimen is not necessarily a simple choice.[207] Given the lack of a clear definition, we use the term adherence. Table 30 summarizes included studies on adherence.

The majority of RCTs that reported adherence stated a rate between 90 percent and 100 percent. Nineteen published studies, examining 18 RCTs, reported levels of adherence.[25,30,35,59,79,85,88-90,128,136,193,194,208-213] Most, however, contained only minimal information, and many did not stratify by treatment. Furthermore, they provided little or no information on the methods of assessment. For example, one fair study reported that both treatment arms exhibited 100 percent adherence, but the investigators did not describe their method of determining adherence.[60] Only 10 of 18 RCTs reported adherence rates for different treatment arms;[30,36,79,81,82,87,88,100,117,212] of these, 8 were head-to-head comparisons (Table 30).[30,79,84,88-90,97,102] None of these studies noted a significant difference in adherence.

None of the three effectiveness studies reported on adherence. To what extent results from highly controlled efficacy trials can be extrapolated to effectiveness settings remains unclear.

A retrospective database analysis used the Integrated Healthcare Information Services National Managed Care Benchmark Database to examine adherence levels in 116,090 patients being treated with SSRIs (immediate-release citalopram, escitalopram, fluoxetine, or sertraline) compared with controlled-release paroxetine.[206] Their primary finding was that patients on a controlled-release formula were significantly more compliant than patients on immediate-release formulations. After controlling for baseline covariates (age, sex, insurance type, titration rates, mental health specialty care, diagnoses, and comorbidity), patients initiating an immediate-release SSRI were 13.6 percent less likely to be adherent than patients on controlled-release paroxetine ($P < 0.0001$).

Table 30. Head-to-head trials reporting adherence to second-generation antidepressants

Study	N Drugs and Dose Duration	Rate of Adherence	Quality Rating
Coleman et al., 1999[88]	364 Bupropion SR 150-400 mg/d Sertraline 50-200 mg/d Placebo 8 weeks	Tablet: Bupropion SR 96% Sertraline 97% Placebo 96% Capsule: Bupropion SR 98% Sertraline 98% Placebo 98%	Fair
Coleman et al., 2001[79]	456 Bupropion SR 150-400 mg/d Fluoxetine 20-60 mg/d Placebo 8 weeks	97% to 99% in all groups	Fair
Croft et al., 1999[89]	360 Bupropion SR 150-400 mg/d Sertraline 50-200 mg/d Placebo 8 weeks	Bupropion SR 98% Sertraline 97% Placebo 98%	Fair
Ekselius et al., 1997[30]	400 Citalopram 20-60 mg/d Sertraline 50-100 mg/d 24 weeks	Citalopram 95% Sertraline 90%	Good
Kavoussi et al., 1997[90]	248 Bupropion SR 100-300 mg/d Sertraline 50-200 mg/d 16 weeks	Bupropion SR 98% Sertraline 99%	Fair
Segraves et al., 2000[102]	248 Bupropion SR 100-300 mg/day Sertraline 50-200 mg/day 16 weeks	Bupropion 98% Sertraline 99%	Fair
Weihs et al., 2000[84]	100 Bupropion SR 100-300 mg/d Paroxetine 10-40 mg/d 6 weeks	Bupropion SR 95% Paroxetine 98%	Good
Weisler et al., 1994[97]	124 Bupropion 225-450 mg/day Trazodone 150-400 mg/day 6 weeks	Bupropion 95% Trazodone 90%	Fair

CR, controlled release; IR, immediate release; SR, sustained release.

Key Question 5: Efficacy, effectiveness, and harms for selected populations

KQ 5. How do the efficacy, effectiveness, or harms of treatment with antidepressants for a depressive syndrome differ for the following subpopulations:

- elderly or very elderly patients;

- other demographic groups (defined by age, ethnic or racial groups, and sex);
- patients with medical comorbidities (e.g., ischemic heart disease, cancer);
- patients with psychiatric and behavioral comorbidities (e.g., substance abuse disorders); and
- patients taking other medications.

Overview

We did not find any studies directly comparing efficacy, effectiveness, and harms of second-generation antidepressants between subgroups and the general population for the treatment of depression syndromes (depressive disorders), which include MDD, dysthymia, and subsyndromal depression including minor depression. Numerous studies, however, conducted subgroup analyses or used subgroups as the study population.

Overall, we included 44 studies in addressing this key question: 36 RCTs, 3 pooled analyses, 2 open-label medication trials (1 randomized), 2 observational studies, and 1 systematic review. We focused on groups defined by age, sex, race or ethnicity, and comorbidities (which included HIV/AIDS, alcohol and substance abuse, Alzheimer's disease or dementia, cardiovascular disease, dialysis, and stroke). These results provide indirect evidence for KQ 5. All studies were rated as fair quality unless otherwise noted.

We present key points and detailed analyses below for the population groups noted above. Details about included studies are presented in Table 31 (listed alphabetically by author within each subgroup). Strength of evidence is moderate for comparative efficacy and effectiveness for age and low for sex, race or ethnicity, and comorbidities; it is low for harms for age.

Age: Key Points

No studies directly compared the efficacy of second-generation antidepressants between either the elderly (60 to 79 years of age) or the very elderly (80 years of age or older) and the general population. Twelve head-to-head efficacy trials[31,34,39,44,45,49,52,55,73,84,94,114,214,215] and two meta-analyses (pooled analyses of original data)[216,217] provide mixed evidence on differences in efficacy in the elderly or very elderly treated with second-generation antidepressants. Comparisons were available for less than one-fourth of the potential comparisons between the 12 second-generation antidepressants addressed in this report. Only seven head-to-head trials or meta-analyses compared the efficacy of any non-SSRI second-generation antidepressant with any other second-generation antidepressant.[55,73,84,94,215-217]

Subgroup analyses of two effectiveness studies provide mixed evidence on differences in effectiveness of second-generation antidepressants between either the elderly or very elderly and the general population for the treatment of MDD,[49] dysthymia, and minor depression.[112,113] We did not find any studies directly comparing the harms of second-generation antidepressants between either the elderly or very elderly and the general population. Some findings from randomized controlled trials and observational evidence indicate that very elderly patients might have an increased risk for some rare but potentially serious adverse events such as hyponatremia and weight loss.[167,200]

Table 31. Studies of efficacy, effectiveness, and harms for patient subgroups

Study	Interventions	N	Results	Quality Rating
Age				
Roose et al., 2004[219]	Citalopram vs. placebo	174	No significant difference in response/remission except in high severity group	Fair
Rocca et al., 2005[114]	Citalopram vs. sertraline	138	No significant difference	NA
Allard et al., 2004[55]	Citalopram vs. venlafaxine XR	151	No significant difference	Fair
Burt et al., 2005[218]	Duloxetine vs. placebo	114	Duloxetine was more efficacious (response/remission); no difference in effect in women 40-55 vs. older or younger women	Fair
Kasper et al., 2005[31]	Escitalopram vs. fluoxetine vs. placebo	517	No significant difference in response rates; remission rates lower for fluoxetine than escitalopram	Fair
Cassano et al., 2002[34]	Fluoxetine vs. paroxetine	242	No significant difference	Fair
Schone and Ludwig, 1993[39] Geretsegger et al., 1994[214]	Fluoxetine vs. paroxetine	106	Greater response rate for paroxetine	Fair
Kroenke et al., 2001[49]	Fluoxetine vs. paroxetine vs. sertraline	573	No significant difference	Fair
Devanand et al., 2005[100]	Fluoxetine vs. placebo	90	No difference in response rates and quality of life	Good
Goldstein et al., 1997[167]	Fluoxetine vs. placebo	671	Higher weight loss with fluoxetine in older patients	Fair
Tollefson et al., 1993[222] Tollefson, et al., 1995[223] Small et al., 1996[224]	Fluoxetine vs. placebo	671	Significantly greater response with fluoxetine; current physical illness not associated with response	Fair
Newhouse et al., 2000[44] Finkel et al., 1999[45]	Fluoxetine vs. sertraline	236	Overall similar efficacy, although sertraline patients experienced greater cognitive improvement and greater response among people over 70 years of age	Fair
Rossini et al., 2005[52]	Fluvoxamine vs. sertraline	93	No significant difference in response rates	Fair
Halikas et al., 1995[94]	Mirtazapine vs. trazodone vs. placebo	150	No significant difference	Fair
Weihs et al., 2000[84]	Paroxetine vs. bupropion SR	100	No differences	Good
Schatzberg et al., 2002[73]	Paroxetine vs. mirtazapine	255	Greater early efficacy for mirtazapine; similar number of CGI responders at end of continuation phase	Fair
Rapaport et al., 2003[213]	Paroxetine (CR and IR) vs. placebo	319	Significantly more responders and remitters for paroxetine (CR and IR formulations) than for placebo	Fair
Barrett et al., 2001[113] Williams et al., 2000[112]	Paroxetine vs. placebo vs. behavioral therapy	656	In patients older than 60 years, significantly greater improvement in symptom scores for paroxetine than for placebo; in patients younger than 60 years, no difference	Fair

CGI, Clinical Global Impressions; CR, controlled release; HRT, hormone replacement therapy; IR, immediate release; MI, myocardial infarction; NA, not applicable; SR, slow release; XR, extended release.

Table 31. Studies of efficacy, effectiveness, and harms for patient subgroups (continued)

Study	Interventions	N	Results	Quality Rating
Schneider et al., 2003[220] Sheikh et al., 2004[221]	Sertraline vs. placebo	752	Significantly more responders in sertraline group both with and without comorbid medical illness	Fair
Wilson et al., 2003[137]	Sertraline vs. placebo	113	No difference in prevention of depression; sertraline associated with longer time to recurrence	Fair
Oslin et al., 2003[215]	Sertraline vs. venlafaxine	52	No significant difference in efficacy; tolerability was lower for venlafaxine	Poor
Kirby et al., 2002[200]	SSRIs use vs. venlafaxine	199	Higher rate of hyponatremia in patients on SSRIs and venlafaxine	Fair
Entsuah et al., 2001[217] Thase et al., 2005[216]	Venlafaxine (IR and XR) vs. SSRIs vs. placebo	2,045	Venlafaxine response not affected by age or sex; SSRI response poorer in older women; similar efficacy of venlafaxine and SSRIs except in older women, but HRT appears to eliminate the difference	Fair
Sex				
Kennedy et al., 2000[188]	Paroxetine vs. sertraline vs. venlafaxine vs. moclobemide	107	Sex difference in impairment in drive/desire; rates of dysfunction in men similar in all treatments; in women, greater levels of dysfunction with sertraline and paroxetine; favorable drug response associated with less dysfunction	Fair
Thase et al., 2005[216] and Enstuah et al., 2001[217]	SSRIs vs. venlafaxine XR vs. placebo	2,045	Venlafaxine response not affected by age or sex; SSRI response poorer in older women; similar efficacy of venlafaxine and SSRIs, except in older women, but HRT appears to eliminate the difference	Fair
Ethnicity				
Wagner et al., 1998[225]	Fluoxetine vs. placebo	118	Ethnicity not associated with side effects; whites had a higher response rate, Latinos a higher drop-out rate	Poor
Comorbidities				
HIV/AIDS				
Rabkin et al., 1999[228]	Fluoxetine vs. placebo	120	No difference in depressed HIV/AIDS patients	Fair
Wagner et al., 1998[225]	Fluoxetine vs. placebo	118	Ethnicity not associated with side effects; whites had a higher response rate, Latinos a higher drop-out rate	Poor
Rabkin et al., 2004[227]	Fluoxetine vs. testosterone vs. placebo	123	No difference in depressed HIV/AIDS patients	Fair
Ferrando et al., 1997[226]	Sertraline vs. paroxetine vs. fluoxetine	33	Completers (all treatment groups) experienced improvements in affective and somatic symptoms (many of which were attributed to HIV rather than depression)	Poor

Table 31. Studies of efficacy, effectiveness, and harms for patient subgroups (continued)

Study	Interventions	N	Results	Quality Rating
Alcohol				
Hernandez-Avila et al., 2004[208]	Nefazadone vs. placebo	41	No significant differences	Fair
Gual et al., 2003[229]	Sertraline vs. placebo	83	No significant differences	Fair
Moak et al., 2003[210]	Sertraline vs. placebo	82	Greater depression improvement in females treated with sertraline; less drinking associated with greater depression improvement	Fair
Alzheimer's disease/dementia				
Nyth et al., 1992[211]	Citalopram vs. placebo	149	Significantly greater improvement with citalopram	Poor
Lyketsos et al., 2003[230]	Sertraline vs. placebo	44	Sertraline associated with greater response	Fair
Magai et al., 2000[231]	Sertraline vs. placebo	31	No significant difference	Fair
Breast cancer				
Roscoe et al., 2005[232]	Paroxetine vs. placebo	94	Paroxetine associated with greater depression response	Poor
Cardiovascular disease				
Strik et al., 2000[235]	Fluoxetine vs. placebo	54	Significantly greater response with fluoxetine	Good
Krishnan et al., 2001[209]	Sertraline	220	Vascular comorbidity not associated with more adverse events or premature discontinuation	Fair
Glassman et al., 2002[234]	Sertraline vs. placebo	369	Significantly greater response with sertraline	Fair
Bush et al., 2005[233]	SSRIs	NR	SSRIs improve depression in post-MI patients	Fair
Dialysis				
Blumenfield et al., 1997[236]	Fluoxetine vs. placebo	14	No significant difference	Fair
Stroke				
Andersen et al., 1994[237]	Citalopram vs. placebo	285	Significantly more improvement with citalopram	Fair
Murray et al., 2005[238]	Sertraline vs. placebo	123	No difference in response; greater improvements in quality of life with sertraline	Fair
Substance abuse				
Petrakis et al., 1998[212]	Fluoxetine vs. placebo	44	No difference in depressed opioid addicts	Fair
Schmitz et al., 2001[239]	Fluoxetine vs. placebo	68	No difference in depressed cocaine abusers	Poor

Age: Detailed Analysis

Head-to-head evidence. We identified 12 head-to-head RCTs (14 articles) in elderly or very elderly patients.[31,34,39,44,45,49,52,55,73,84,94,114,214,215] We also identified one set of meta-analyses of original data from eight RCTs.[216,217] These trials evaluated numerous treatment comparisons: first, intra-SSRI comparisons (citalopram vs. sertraline, escitalopram vs. fluoxetine vs. placebo, fluoxetine vs. paroxetine, fluoxetine vs. sertraline, fluvoxamine vs. sertraline, and paroxetine vs. sertraline); second, SSRI vs. SNRI comparisons (citalopram vs. venlafaxine XR, paroxetine vs. mirtazapine, sertraline vs. venlafaxine IR, SSRIs vs. venlafaxine IR or XR vs. placebo); third,

SSRIs vs. other second-generation antidepressants (paroxetine vs. bupropion SR); and fourth, SNRIs vs. other second-generation antidepressants (mirtazapine vs. trazodone vs. placebo).

SSRIs vs. SSRIs. *Citalopram vs. sertraline.* One randomized trial evaluated citalopram and sertraline in the treatment of 138 nondemented elderly patients with minor depressive disorder and subsyndromal depressive symptomatology.[114] Although this trial does not meet eligibility criteria because of the study design (because of flawed randomization, it is essentially a nonrandomized trial) and therefore is not assigned a quality rating, we included it here because it is the only evidence pertaining to a comparison of these two SSRIs; the trial also had high loss to followup (27.5 percent). Both treatments improved depressive symptoms (as measured by the HAM-D); HAM-D remission rates were similar for citalopram and sertraline at the end of the study (53 percent and 42 percent, $P = 0.25$). Similar improvements were seen in Global Assessment of Function (GAF) and cognitive scores.

Escitalopram vs. fluoxetine vs. placebo. One 8-week study compared escitalopram, fluoxetine, and placebo in 517 participants older than 65 years of age (mean age in each treatment group, 75 years).[31] Outcome measures included the MADRS and the CGI-S. Patients on escitalopram experienced greater improvement than those on fluoxetine in MADRS score at week 8 (using an LOCF analysis) ($P < 0.01$); however, the patients treated with escitalopram and with placebo did not differ significantly. Escitalopram, placebo, and fluoxetine MADRS response rates were similar (46 percent, 47 percent, 37 percent, respectively, P = not significant).

In addition, MADRS remission rates were similar for escitalopram and placebo (40 percent and 42 percent), but for fluoxetine vs. placebo the difference was significant (30 percent vs. 42 percent, $P = 0.05$). Escitalopram- and fluoxetine-treated patients experienced significantly more nausea than placebo-treated patients ($P < 0.01$).

Fluoxetine vs. paroxetine. We identified three randomized trials (four articles) in populations older then 60 years of age.[34,39,49,214] One 6-week trial (two publications) compared fluoxetine (20-60 mg/day) and paroxetine (20-40 mg/day) in 106 depressed patients ages 61 to 85 years (mean age 74 years).[39,214] In this trial, patients treated with paroxetine achieved statistically significantly higher HAM-D response rates than patients in the fluoxetine group ($P = 0.03$).[39,214] No significant differences were seen in overall adverse events.

By contrast, an effectiveness trial conducted in patients older than 18 years of age with major depression, dysthymia, or minor depression did not detect any differences in the effectiveness of fluoxetine and paroxetine.[49] Both treatment groups showed significant improvements in depression and other health-related quality of life domains (social function, work function, physical function) with no significant differences between study groups. It also produced no interactions between treatment groups and age (≥ 60 years of age vs. younger).

An Italian study lasting 1 year enrolled 242 patients to compare the effects of fluoxetine (20-60 mg/day) and paroxetine (20-40 mg/day) on depressive symptoms, mood, and cognitive function in nondemented persons 65 years or older.[34] Treatment groups did not differ significantly at study endpoint in CGI scores. Although there are no statistically significant differences in outcome measures, this finding does not conclusively demonstrate noninferiority.[109] Severe adverse events were significantly more common in the fluoxetine group than the paroxetine group (22 vs. 9 events; $P < 0.002$).

Fluoxetine vs. sertraline. One 12-week study (two articles) comparing sertraline (50-100 mg/day) and fluoxetine (20-40 mg/day) in 236 participants ages 60 years and older provides evidence of the comparable efficacy of these drugs.[44,45] Outcome measures included MADRS, HAM-D, quality of life (Quality of Life Enjoyment and Satisfaction Questionnaire), and cognitive assessments (Shopping List Task [SLT], MMSE, and Digital Symbol Substitution Test [DSST]). Patients treated with these drugs did not differ significantly on primary outcome measures (MADRS, HAM-D). HAM-D response rates (sertraline, 73 percent; fluoxetine, 71 percent) and HAM-D remission rates (sertraline, 45 percent; fluoxetine, 46 percent) were similar. Adverse event rates were similar in the two treatment groups. Quality of life and other patient-rated measures were also similar for both treatment groups at endpoint. Sertraline-treated patients showed greater cognitive improvement than patients on fluoxetine on the DSST at endpoint ($P = 0.037$). A subgroup analysis of 75 patients 70 years of age or older demonstrated a greater response rate for sertraline than for fluoxetine (58.5 percent vs. 42.4 percent, respectively, $P = 0.027$).[45]

A 9-month effectiveness study yielded similar results.[49] The investigators found no differences in effectiveness between fluoxetine and sertraline in patients older than 18 years of age with major depression, dysthymia, or minor depression. Both treatment groups showed significant improvements in depression and other health-related quality of life domains (social function, work function, physical function) with no significant differences between study groups. No interactions between treatment groups and age (≥ 60 years of age vs. younger) were seen.

Fluvoxamine vs. sertraline. A 7-week trial compared fluvoxamine and sertraline for the treatment of major depression in 93 patients 59 years of age and older (mean age for both treatment groups, 68 years).[52] HAM-D response rates favored fluvoxamine over sertraline but did not reach statistical significance (71.8 percent vs. 55.6 percent, $P = 0.12$). Although the difference was not statistically significant, this finding does not conclusively demonstrate noninferiority.[109] A two-way repeated measures analysis of variance revealed a significant time-by-group interaction in HAM-D scores favoring fluvoxamine ($P = 0.007$).

Paroxetine vs. sertraline. A fair effectiveness trial conducted in patients older than 18 years of age with major depression, dysthymia, or minor depression did not detect any differences in the effectiveness of paroxetine and sertraline.[49] The investigators found no differences in effectiveness between paroxetine and sertraline in patients 18 years of age and older with major depression, dysthymia, or minor depression. Both treatment groups showed significant improvements in depression and other health-related quality of life domains (social function, work function, physical function) with no significant differences between study groups. No interactions between treatment groups and age (≥ 60 years of age vs. younger) were seen.

SSRIs vs. SNRIs. *Citalopram vs. venlafaxine XR.* A European 6-month study compared citalopram with venlafaxine XR for the treatment of depression in 151 elderly outpatients (mean age, 73 years).[55] The investigators found no statistically significant differences at study endpoint in any outcome measures (MADRS, CGI-S, CGI-I). MADRS remission rates were 23 percent for citalopram and 19 percent for venlafaxine (P = not reported). Both treatment groups reached a 93 percent response rate at week 22 (response defined as a reduction of at least 50 percent in MADRS score). Although outcome measures did not differ significantly, this finding does not conclusively demonstrate noninferiority.[109] More spontaneously reported adverse events were

reported by venlafaxine-treated patients than citalopram-treated patients (62 percent vs. 43 percent, respectively); tremor was more common in the citalopram group than the venlafaxine group, and nausea or vomiting was more common in the venlafaxine group than the citalopram group.

Paroxetine vs. mirtazapine. One study compared paroxetine (20-40mg/day) and mirtazapine (15-45 mg/day) for the treatment of major depression in 255 elderly patients 65 years of age and older; the trial included an acute phase (8 weeks) and a continuation phase (16 weeks).[73] Mirtazapine was associated with significantly more patients who responded (50 percent reduction in HAM-D scores) at day 14 and patients in remission at day 42 (27.8 percent and 31.0 percent, respectively) than paroxetine (13.3 percent and 19.2 percent, respectively; $P = 0.005$ and $P = 0.044$, respectively). The median time to achieving response was significantly shorter for mirtazapine than for paroxetine (26 days vs. 40 days, respectively). At study endpoint, the number of CGI responders was similar in the mirtazapine and paroxetine treatment groups (64 percent and 56.7 percent, respectively, $P = 0.267$). Significantly more mirtazapine-treated patients reported dry mouth and weight gain ($P < 0.05$). Paroxetine-treated patients reported a significantly higher rate of nausea, tremor, and flatulence ($P < 0.05$).

Sertraline vs. venlafaxine IR. One 10-week randomized trial compared sertraline (up to 100 mg/day) and venlafaxine IR (up to 150 mg/day) among 52 nursing home residents (61 to 99 years of age).[215] We graded the quality of this study as poor for efficacy because of high loss to followup (44 percent), but we note it here because it is the only study comparing these two agents and because the high loss to followup may be expected in this population (elderly nursing home residents) and may not be a reflection of the quality of the study. The investigators reported a significantly higher rate of loss to followup among venlafaxine- than sertraline-treated patients (63 percent vs. 24 percent). Venlafaxine-treated patients had a significantly higher rate of withdrawal because of severe adverse events ($P = 0.022$) and withdrawal because of severe adverse events or side effects ($P = 0.005$) than did the sertraline-treated patients.

SSRIs vs. venlafaxine (IR or XR) vs. placebo. In one study, investigators pooled data from eight randomized trials of venlafaxine IR (75-375 mg/day) or venlafaxine XR (75-225 mg/day), one of several SSRIs (fluoxetine, 20-80 mg/day; fluvoxamine, 100-200 mg/day; paroxetine, 20-40 mg/day), or placebo in the treatment of depression.[216,217] This study was not based on a systematic literature search, so results must be viewed cautiously. The trials varied in length (6 weeks [three studies], 8 weeks [four studies], or 12 weeks [one study]) and included either outpatients (seven studies) or inpatients (one study). Four of the outpatient trials had a placebo arm. For venlafaxine-treated patients, neither age (< 50 or ≥ 50 years of age) nor sex affected remission rates.[216] Among patients treated with SSRIs, however, a significant interaction was observed between treatment and sex ($P = 0.004$): older women had a poorer SSRI response (response rate: 28 percent) than younger women (response rate: 36 percent) and both older and younger men (response rates: 35 percent and 36 percent, respectively). Remission rates for older women treated with venlafaxine (48 percent) were higher than remission rates for older women treated with SSRIs (28 percent, $P = 0.0004$). Hormone replacement therapy appeared to eliminate these differences. Additional analyses of age subgroups (≤ 40, 41-54, 55-64, and ≥ 65 years of age) and sex subgroups revealed that no significant age-by-treatment, sex-by-treatment, or age-by-sex-by-treatment interactions occurred; men and women of different ages within each

treatment group had similar rates of remission, response, and absence of depressed mood.[217] Among patients over 40 years of age, the rates of adverse events were similar between the treatment groups, although venlafaxine-treated patients 55 to 64 years of age reported significantly more nausea than placebo ($P \leq 0.003$), and placebo patients 41 to 54 years of age reported significantly more headache than venlafaxine ($P \leq 0.01$) .

A fair retrospective cohort study reported that hyponatremia in elderly inpatients (mean age 74 years) was significantly more common in patients treated with SSRIs or venlafaxine than in controls not on these drugs (OR, 3.5; 95% CI, 1.4-8.9).[200] Otherwise, evidence from controlled trials and observational studies is insufficient to conclude for or against an increased risk of hyponatremia in patients treated with SSRIs.

SSRIs vs. other second-generation antidepressants. *Paroxetine vs. bupropion SR.* One good-quality RCT examined the efficacy of paroxetine and bupropion SR over 6 weeks in 100 outpatients ages 60 years or older (range 60-88 years).[84] The majority of patients were white (paroxetine, 90 percent; bupropion SR, 98 percent), female (paroxetine, 60 percent; bupropion SR, 54 percent), and did not use antidepressants for the current episode before enrollment (paroxetine, 88 percent; bupropion SR, 83 percent). Statistical analysis used an LOCF approach. The overall loss to followup was 16 percent, with no significant difference between treatment groups. Overall adverse events were similar in the two treatment groups. Efficacy according to any outcome measure did not differ significantly between treatment groups. Response rates ($\geq$ 50 percent reduction in HAM-D scores) were similar in both groups (paroxetine, 77 percent; bupropion SR, 71 percent).

SNRIs vs. other second-generation antidepressants. *Mirtazapine vs. trazodone vs. placebo.* One study compared mirtazapine with trazodone in patients with MDD older than 55 years of age.[94] Efficacy outcome measures in this trial favored mirtazapine, but differences did not reach statistical significance. Although outcome measures did not differ significantly, this finding does not conclusively demonstrate noninferiority.[109] More mirtazapine-treated patients discontinued treatment than did those on both trazodone and placebo. Both treatments were associated with more somnolence and dry mouth than placebo ($P \leq 0.05$); trazodone treatment was associated with significantly more dizziness and blurred vision compared to placebo ($P \leq 0.05$).

Placebo-controlled evidence. Two studies (four articles) provide evidence of the general efficacy of paroxetine (CR and IR formulations) and fluoxetine in the treatment of elderly patients with depression.[213,222-224] A good-quality trial evaluated the efficacy of fluoxetine for treating patients 60 years of age and older with dysthymia over 12 weeks.[100] ITT results indicated that fluoxetine had limited efficacy. Response rates on the HAM-D favored fluoxetine over placebo, but the two groups did not differ significantly (27.3 percent vs. 19.6 percent; $P < 0.4$). One study of fluoxetine vs. placebo (n = 671 patients older than 60 years of age) recorded a significant weight loss for fluoxetine compared with placebo.[167]

A study of citalopram vs. placebo yielded no significant difference between the two treatment groups in response or remission rates except in the high severity group ($P = 0.04$).[219] One sertraline trial (two publications) demonstrated that sertraline was superior to placebo in elderly patients with late-life depression with and without comorbid medical illness (HAM-D and CGI-I response rates).[220,221] Another trial reported that sertraline was not superior to placebo

in the prevention of recurrence; however, patients on sertraline experienced a longer time to recurrence than did patients on placebo (92 weeks and 48 weeks, respectively).[137]

One large, primary-care-based effectiveness study (two publications) randomized 656 patients with dysthymia or minor depression to 11 weeks of paroxetine, placebo, or behavioral therapy.[112,113] Participants were stratified into patients 60 years and older (n = 415) and patients younger than 60 years (n = 241) for ITT analysis.

In the 60 or older subgroup, paroxetine-treated patients showed a greater change in HSCL-D-20 scores than placebo-treated patients ($P = 0.004$).[112] Effects were similar for patients with dysthymia and minor depression. For older dysthymia patients with high or intermediate baseline functioning scores, paroxetine significantly improved mental health functioning compared with placebo. Overall, however, improvements of mental health functioning were not statistically significantly different between dysthymia patients receiving paroxetine and those receiving placebo.

Among the younger patients, treatment groups did not differ significantly on the HSCL-D scale.[113] For dysthymia only, the remission rate of patients with at least 4 weeks of treatment was significantly higher in the paroxetine group than in the placebo group (80 percent vs. 44 percent; $P = 0.008$). Paroxetine was not more efficacious than placebo in patients with minor depression.

One 9-week pooled analysis trial evaluated the efficacy of duloxetine in 114 depressed women ages 40 to 55 and women younger than 40 or older than 55.[218] In women 40 to 55, both response rates (58.2 percent vs. 32.2 percent, $P = 0.003$) and remission rates (34.6 percent vs. 18.6 percent, $P = 0.027$) were significantly higher in the women receiving duloxetine than in women receiving placebo. The magnitude of treatment effect was similar in women 40 to 55 years of age and older women (more than 55 years of age) and younger women (less than 40 years of age).

Sex: Key Points

Two head-to-head comparisons (one observational trial[188] and one pooled analysis[216,217]) were available for the potential comparisons between the 12 second-generation antidepressants addressed in this report. The effects of second-generation antidepressants do not appear to differ by sex.[188,216,217] Some differences may exist in the frequency of some adverse events (sexual side effects,[188] headache,[216,217] and nausea).[216,217] In both cases, the strength of the evidence is low.[188]

No placebo-controlled trials are available on the efficacy, effectiveness, and harms of second-generation antidepressants in men and women.

Sex: Detailed Analysis

Head-to-head evidence. One head-to-head observational study and one pooled analysis compared men and women.[188,216,217] The pooled analyses do not provide evidence of any differences in efficacy, although some difference was observed in adverse events. Men experienced higher rates of headaches with venlafaxine than with placebo, and women experienced higher rates of nausea with venlafaxine than with placebo.[216,217] Observational evidence also suggests that men and women may experience differences in sexual side effects.[188] The strength of the evidence is low.

SSRIs vs. SNRIs. One 14-week trial of paroxetine (mean dose 30.7 mg/day), sertraline (99.0 mg/day), venlafaxine (151.6 mg/day), and moclobemide (485 mg/day) evaluated disturbances in sexual drive/desire and arousal/orgasm in depressed patients who completed 8 weeks of the study.[188] Men reported greater impairment in drive/desire than women ($P < 0.05$). Men and women did not differ significantly on the arousal/orgasm scale ($P = 0.21$). Rates of dysfunction in all treatment groups were similar for men; among women, sertraline and paroxetine appeared to be associated with greater dysfunction. All drugs appeared to be equally effective in reducing depressive symptoms (main effect for time, $P < 0.001$); a favorable drug response was associated with less sexual dysfunction.

SSRIs vs. venlafaxine (IR or XR) vs. placebo. As described above (head-to-head evidence for age), data were pooled from eight randomized trials of venlafaxine IR (75-375 mg/day) or venlafaxine XR (75-225 mg/day), one of several SSRIs (fluoxetine, 20-80 mg/day; fluvoxamine, 100-200 mg/day; paroxetine, 20-40 mg/day), or placebo in the treatment of depression.[216,217] Remission rates for venlafaxine-treated patients were not affected by sex.[216] Among patients treated with SSRIs, however, a significant interaction was observed between treatment and sex ($P = 0.04$): older women had a poorer SSRI response (28 percent) than younger women (36 percent) and than both older and younger men (35 percent and 36 percent, respectively). Additional analyses of the age (≤ 40, 41-54, 55-64, and ≥ 65) and sex subgroups revealed no significant sex-by-treatment or age-by-sex interactions; men and women of different ages within each treatment group had similar rates of remission, response, and absence of depressed mood.[217]

Placebo-controlled evidence. We did not identify any placebo-controlled trials on the efficacy or harms of second-generation antidepressants in men and women.

Race or Ethnicity: Key Points

One placebo-controlled efficacy study compared different racial or ethnic groups.[225] Fluoxetine and placebo did not differ significantly in outcomes, but this study may not have been powered to detect a significant difference; it was rated poor quality because it lacked an ITT analysis (completer analysis only). No studies directly compared the effectiveness and harms of second-generation antidepressants between different races or ethnicities.

Race or Ethnicity: Detailed Analysis

Head-to-head evidence. No head-to-head trials on the efficacy, effectiveness, or harms of second-generation antidepressants compared different racial or ethnic groups.

Placebo-controlled evidence. One trial evaluated the efficacy of fluoxetine vs. placebo in the treatment of patients with comorbid HIV/AIDS.[225] We included it even though we rated it as poor for efficacy (no ITT analysis) because it is the only trial identified in the literature search that examined race or ethnicity.

A total of 118 patients were randomized to 8 weeks of treatment with either fluoxetine or placebo. Of all participants, 67 percent were white, 19 percent black, and 14 percent Latino; only 1.7 percent (n = 2) were female. Loss to followup was significantly greater among Latinos (53 percent) than blacks (14 percent) and whites (28 percent) ($P < 0.05$). Ethnicity was not

associated with the total number of treatment side effects or dosage. Response rates among subjects who completed the study were higher in the fluoxetine group (white, 84 percent; black, 50 percent; Latino, 67 percent) than the placebo group (white, 43 percent; black, 36 percent; Latino, 80 percent). The differences were not significant; however, this may be because of the small sample size, particularly in the Latino group.

Comorbidities: Key Points

We found no studies directly comparing the efficacy, effectiveness, and harms of second-generation antidepressants between depressed patients with comorbidities and the general population.

One poor-quality head-to-head study examined the efficacy of treatment with second-generation antidepressants for patients with MDD and comorbid HIV/AIDS.[226]

Eighteen placebo-controlled trials of varying quality[208-212,225,227-232,234-239] and one systematic review[233] evaluated second-generation antidepressants in patients with various comorbid conditions. Some of these suggested that second-generation antidepressants may not be efficacious for depressed patients with comorbidities.[208,210,212,225,227-229,231,236-239] However, many of the studies may not have been powered to detect a difference between active treatment and placebo.

Comorbidities: Detailed Analysis

HIV/AIDS. One poor-quality head-to-head study compared the efficacy and tolerability of fluoxetine, paroxetine, or sertraline in depressed individuals with HIV.[226] This 6-week open-label trial evaluated 33 depressed HIV-positive men and women. This trial was rated as poor because it included a completer-only analysis (no ITT analysis); it has been included here because it is the only head-to-head trial in patients with depression and various comorbidities. The overall clinical response rate for completers (n = 24) was 83 percent (fluoxetine, 90 percent; paroxetine, 86 percent; sertraline, 71 percent). Overall, and in each treatment group, significant reductions were seen in both affective and somatic symptoms (as measured by the HAM-D, BDI, HAM-D affective subscale, BDI cognitive subscale, HAM-D vegetative subscale, and BDI somatic subscale scores among completers), including somatic symptoms that were attributed to HIV rather than depression.

Two placebo-controlled studies evaluated the efficacy of fluoxetine vs. placebo in the treatment of patients with depression and comorbid HIV/AIDS.[227,228] The first study, a 12-week randomized trial, compared fluoxetine and placebo in the treatment of depression in patients with HIV/AIDS.[228] The second trial, a 12-week, randomized trial compared fluoxetine, testosterone, and placebo in the treatment of depression in patients with HIV/AIDS.[227] In both studies, fluoxetine and placebo response rates (57 percent vs. 41 percent[228] and 54 percent vs. 44 percent[227]) did not differ significantly. However, these studies may not have been powered to detect a statistically significant difference.

A third, 8-week, placebo-controlled trial evaluated the efficacy of fluoxetine vs. placebo in the treatment of patients with comorbid HIV/AIDS (described for race and ethnicity).[225] We rated it as poor because it had no ITT analysis; however, we included it here because of the very limited evidence on this topic. Response rates among subjects who completed the study were higher in the fluoxetine group (white, 84 percent; black, 50 percent; Latino, 67 percent) than in

the placebo group (white, 43 percent; black, 36 percent; Latino, 80 percent). The differences were not significant; however, this may be because of the small sample size, particularly in the Latino group.

Alcohol abuse. One randomized trial compared nefazadone and placebo in the treatment of depressed patients with comorbid alcohol dependence over a period of 10 weeks.[208] Nefazadone was superior to placebo, as measured by improvement in depression on the HAM-D from intake to study endpoint (mean change in HAM-D score for nefazadone vs. placebo: -12.25 vs. -12.55, $P = 0.51$), the difference did not achieve statistical significance, perhaps because the study was underpowered to do so.

Two randomized trials compared sertraline and placebo in the treatment of patients with depression and alcoholism.[210,229] Results suggested that, in some subgroups, sertraline was superior to placebo (i.e., in women), but overall the two treatment groups did not differ significantly. A 24-week study compared sertraline (50-150 mg/day) and placebo in recently detoxified alcohol-dependent patients with current depressive symptoms.[229] Response (≥ 50 percent decrease in MADRS) was slightly higher in sertraline-treated patients (44 percent) than in placebo-treated patients (39 percent). Both groups experienced significant improvements in HAM-D and MADRS scores during the study, although the two groups did not differ significantly. Adverse event rates were similar for the two treatment groups. A 12-week trial showed similar results.[210] In women, treatment with sertraline was associated with less depression at the end of treatment than those receiving placebo. Less drinking during the study was associated with improved depression outcomes.

Alzheimer's disease or dementia. Two randomized trials compared sertraline and placebo for patients with depression and comorbid Alzheimer's disease.[230,231] An 8-week trial of late-stage Alzheimer's disease failed to demonstrate a statistically significant difference between sertraline and placebo; 47 percent and 36 percent, respectively, achieved at least a 50 percent improvement in the Cornell Score for Depression in Dementia (CSDD) and 35 percent and 50 percent, respectively, achieved at least a 50 percent improvement in the Gestalt Depression Scale. However, this study may not have been powered to detect statistically significant differences.[231] A fair 12-week trial demonstrated that sertraline was statistically significantly superior to placebo, as measured by both the CSDD ($P = 0.002$) and the Hamilton Depression Rating Scale (HDRS) ($P = 0.01$).[230] More patients treated with sertraline responded to treatment (full responders, 38 percent; partial responders, 46 percent) than did patients treated with placebo (full responders, 20 percent; partial responders, 15 percent) ($P = 0.007$).

One poor-quality randomized trial compared citalopram and placebo for patients 65 years of age and older with depression and comorbid mild to moderate dementia.[211] We rated this trial poor because it appeared to be a completer-analysis only and had high loss to followup. In the efficacy analysis, which includes only those patients who completed the trial, the mean total HAM-D score at endpoint ($P < 0.05$) and improvement in HAM-D total score at endpoint ($P < 0.01$) were statistically significantly better for patients treated with citalopram than those receiving placebo; similar results were seen with the CGI-S. Significantly more citalopram-treated patients than placebo-treated patients improved (score of 1 or 2 on the CGI-I) (60 percent vs. 24 percent, $P < 0.001$).

Breast cancer. One randomized trial compared paroxetine and placebo for patients with breast cancer who were receiving at least four cycles of chemotherapy to evaluate whether the use of an antidepressant can alleviate symptoms of depression and reduce fatigue.[232] We rated it as poor because it appeared to be a completer-analysis only and the length of the study was not adequately described. Paroxetine was more effective in reducing depression during chemotherapy, as measured by the Center for Epidemiological Studies of Depression (CES-D) ($P = 0.006$); mean (standard deviation [SD]) scores at cycle 4 for paroxetine and placebo were 8.8 (1.11) and 12.6 (1.24), respectively. However, paroxetine and placebo did not differ significantly on all four fatigue scales.

Cardiovascular disease. AHRQ sponsored a systematic review of postmyocardial infarction (post-MI) depression that we graded fair overall. The authors concluded that SSRIs improve depression in post-MI patients.[233]

We also identified three studies evaluating second-generation antidepressants in the treatment of depression in patients with MI or angina;[209,234,235] two had been included in the AHRQ report.[234,235] The first, a 24-week randomized trial, evaluated sertraline vs. placebo for treating depression in patients with acute MI or unstable angina.[234] The second, a good-quality 25-week randomized trial, evaluated fluoxetine vs. placebo in the treatment of depression after a first MI.[235] In both trials, active treatment was associated with a significantly greater response rate than placebo (sertraline, 67 percent; placebo, 53 percent; $P = 0.01$;[234] fluoxetine, 48 percent; placebo, 26 percent; $P = 0.05$[235]).

The third study was a pooled analysis of two randomized trials of sertraline (sertraline vs. fluoxetine and sertraline vs. nortriptyline); the article evaluated sertraline only in patients older than 60 with vascular disease (fluoxetine and nortriptyline results are not reported).[209] Newhouse et al. reported the results of the sertraline vs. fluoxetine comparison in all patients (not limited to those with cardiovascular disease); this article is described in detail for KQ 5, subsection Age.[44] Results for sertraline vs. nortriptyline were reported by Bondareff et al.;[240] this article was excluded because the comparison (sertraline vs. nortyriptyline) is outside of the scope of interest of this report.

The former analysis categorized patients into one of three groups: patients with a current diagnosis of hypertension but no other past or present cardiovascular disease, patients reporting a current or past history of cardiovascular illness but excluding hypertension, and patients with no hypertension and no other comorbid vascular illness. Sertraline was safe, well tolerated, and effective as an antidepressant in elderly patients suffering from hypertension and other forms of vascular comorbidity. Rates of response (measured by the HAM-D and the CGI-I) were similar in sertraline-treated patients in all three vascular illness categories.

Dialysis. We identified one randomized trial of fluoxetine vs. placebo in depressed patients on dialysis (N = 14).[236] Patients treated with fluoxetine had slightly greater improvement than patients treated with placebo, as measured by the BDI (-9.57 vs. -8.8, $P = 0.91$), the Brief Symptom Inventory (Depression Scale) (-4.43 vs. -3.2, $P = 0.88$), the HAM-D (-9.0 vs. -7.5, $P = 0.72$), and the MADRS (-11.14 vs. -6.67, $P = 0.45$). Although the differences were not statistically significant at study endpoint, this may be attributable to the small sample size. No patients discontinued because of side effects; no side effects were judged to be severe.

Stroke. Two studies evaluated the efficacy of citalopram and sertraline in the treatment of patients with poststroke depression.[237,238] One 6-week randomized trial evaluated the efficacy of citalopram vs. placebo in poststroke depression.[237] A 26-week trial evaluated the efficacy of sertraline vs. placebo in the treatment of minor depression and less severe depression in stroke patients.[238] Citalopram was associated with significantly greater improvements in depression than placebo on the HAM-D; mean (SD) improvements for citalopram vs. placebo were 8.0 (6.0) vs. 7.2 (5.8), respectively.[237] Sertraline and placebo did not differ significantly in response rates (week 6: 56 percent vs. 46 percent, respectively; week 26: 76 percent vs. 78 percent, respectively) or week 6 or week 26 remission rates (week 6: 59 percent vs. 51 percent, respectively; week 26: 81 percent vs. 87 percent, respectively).[238] However, at week 26, sertraline was associated with greater improvements in quality of life than placebo (effect size not reported, $P < 0.05$).

Substance abuse. Two studies evaluated the efficacy of fluoxetine in the treatment of patients with depression and comorbid substance abuse (persons with methadone-maintained opioid addiction or cocaine dependence); overall, fluoxetine and placebo did not differ significantly in reducing depressive symptoms.[212,239]

One randomized 12-week trial evaluated fluoxetine vs. placebo in the treatment of depression in methadone-maintained opioid addicts.[212] Among the entire sample (N = 44), BDI (mean decrease for fluoxetine vs. placebo -8.0 vs. -4.7, respectively) and HDRS scores (mean decrease for fluoxetine vs. placebo: -6.0 vs. -7.7, respectively) decreased in both groups, but the treatment groups did not differ significantly. Among those subjects with major depression (n = 31), there were no significant differences in the rate of change of depressive symptoms by treatment group (fluoxetine vs. placebo) over time (BDI: -7.8 vs. -3.4; respectively; HDRS: -5.1 vs. -6.9, respectively). The second study, a poor-quality, 12-week randomized trial, evaluated fluoxetine vs. placebo for treating major depression in cocaine-dependent patients.[239] This trial was rated poor for efficacy (high loss to followup [52.9 percent]) but is included here because of the dearth of evidence on this topic. Fluoxetine and placebo did not differ significantly on the BDI (effect size not reported).

Discussion

General Conclusions

This report provides a comprehensive summary of the comparative efficacy, effectiveness, and harms of 12 second-generation antidepressants for the treatment of major depressive disorder (MDD), dysthymia, and subsyndromal depression. They include bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine in three classes: selective serotonin reuptake inhibitors (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs and SSNRIs), and other second-generation antidepressants. Table 32 briefly summarizes our findings from evidence for all five key questions and their subquestions and notes the strength of evidence in each case.

Most of the relevant trials were conducted in patients with MDD. Therefore, we can draw some conclusions regarding the use of second-generation antidepressants for MDD. Evidence is insufficient, however, to draw firm conclusions about comparative efficacy, effectiveness, and harms of second-generation antidepressants for dysthymia and subsyndromal depression.

For MDD, our findings indicate that the existing evidence does not warrant the choice of one second-generation antidepressant over another based on greater efficacy and effectiveness. We could not find any substantial differences in efficacy and effectiveness for either treating the acute depressive phase or maintaining remission. Furthermore, no differences in efficacy and effectiveness are apparent in subgroups based on age and sex, although evidence within subgroups is more limited.

More than 50 percent of patients treated with second-generation antidepressants for acute-phase depression did not achieve remission, the goal of depression treatment. Almost 40 percent of patients failed to respond, a less rigorous outcome. Currently, the evidence is insufficient to determine patient factors that can reliably predict response or nonresponse to an individual drug.

Although limited evidence indicates that second-generation antidepressants are also similar in efficacy for treating patients who had failed to respond to a first-line agent, a substantial proportion of these patients do not achieve response or remission with second-line treatment. Multiple treatment options, therefore, are required for patients who do not respond to first- or second-line treatment.

Clinically, numerous physical and psychological symptoms accompany depressive disorders. Clinicians sometimes recommend using individual second-generation antidepressants for these problems, assuming differences in efficacy to treat these accompanying symptom clusters. The current evidence does not support the selection of one second-generation antidepressant over another for specific accompanying symptoms. The best comparative evidence suggests no difference in efficacy for anxiety symptoms. For other symptom clusters such as melancholia, psychomotor change, pain, and somatization, the evidence is limited to few comparisons. For other common symptoms, such as fatigue and loss of energy, evidence is lacking.

Table 32. Summary of findings with strength of evidence

Key Question, Disorder, and Outcome of Interest	Strength of Evidence*	Findings†
Key Question 1a. Comparative efficacy and effectiveness of second-generation antidepressants		
Major depressive disorder		
Comparative efficacy	Moderate	Results from direct and indirect comparisons indicate that no substantial differences in efficacy exist among second-generation antidepressants.
Comparative effectiveness	Moderate	Direct evidence from one good and two fair effectiveness studies and indirect evidence from efficacy trials indicate that no substantial differences in effectiveness exist among second-generation antidepressants.
Quality of life	Moderate	Consistent results from 18 mostly fair studies indicate that the efficacy of second-generation antidepressants with respect to quality of life does not differ among drugs.
Onset of action	Moderate	Consistent results from seven fair trials suggest that mirtazapine has a significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline. Whether this difference can be extrapolated to other second-generation antidepressants is unclear. Most other trials do not indicate a faster onset of action of one second-generation antidepressant compared with another.
Dysthymia		
Comparative efficacy	Low	No head-to head evidence exists. Findings from five placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	Low	One fair effectiveness study provides mixed evidence about paroxetine vs. placebo; patients older than 60 showed greater improvement on paroxetine; those younger than 50 did not show any difference.
Quality of life	No evidence	
Onset of action	No evidence	
Subsyndromal depression		
Comparative efficacy	Low	One nonrandomized, open-label trial did not detect any difference between citalopram and sertraline. Findings from two placebo-controlled trials were insufficient to draw conclusions.
Comparative effectiveness	No evidence	
Quality of life	No evidence	
Onset of action	No evidence	
Key Question 1b: Greater efficacy and effectiveness with previously effective medications		
Major depressive disorder	No evidence	
Dysthymia	No evidence	
Subsyndromal depression	No evidence	

*Strength of evidence is based on a modified version of the GRADE system.[20]
†Good, fair, or poor designations relate to quality grades given to each study; see Methods.
RCT, randomized controlled trials; SR, slow release; XR, extended release.

Table 32. Summary of findings with strength of evidence (continued)

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Key Question 2a: Efficacy and effectiveness of second-generation antidepressants for maintaining response or remission (i.e., preventing relapse or recurrence)		
Comparative efficacy	Moderate	Based on findings from three efficacy trials, no significant differences exist between fluoxetine and sertraline, fluvoxamine and sertraline, and trazodone and venlafaxine for preventing relapse or recurrence. Whether this finding can be extrapolated to other second-generation antidepressants is unclear.
Comparative effectiveness	No evidence	
General effectiveness/efficacy	Moderate	Based on findings from 21 placebo-controlled trials, second-generation antidepressants are effective for preventing relapse or recurrence.
Key Question 2b: Efficacy and effectiveness of second-generation antidepressants in managing treatment-resistant depression syndrome or treating recurrent depression		
Managing treatment-resistant depression		
Comparative efficacy	Low	Results from one fair trial support modestly better efficacy for venlafaxine compared with paroxetine.
Comparative effectiveness	Moderate	Results from two effectiveness studies are conflicting. Based on one trial rated good, no significant differences in effectiveness exist among bupropion SR, sertraline, and venlafaxine XR. One fair effectiveness trial found venlafaxine to be modestly superior to citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline.
General effectiveness/efficacy	Low	No placebo-controlled evidence exists. Uncontrolled, open-label evidence supports the general efficacy of second-generation antidepressants.
Treating recurrent depression		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Key Question 3a: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of depression in patients with accompanying symptom clusters		
Anxiety		
Comparative efficacy	Moderate	Results from six head-to-head trials and one placebo-control led trial (all fair quality) suggest that efficacy does not differ substantially for treatment of depression in patients with accompanying anxiety.
Comparative effectiveness	No evidence	
Insomnia		
Comparative efficacy	Low	Evidence from three fair head-to-head studies is insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting insomnia. Results are limited by study design.
Comparative effectiveness	No evidence	

Table 32. Summary of findings with strength of evidence (continued)

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Melancholia		
Comparative efficacy	Low	Evidence from two fair head-to-head studies, one poor head-to-head study, and one fair placebo-controlled trial is insufficient to draw conclusions about treating depression in patients with coexisting insomnia. Results are inconsistent across studies.
Comparative effectiveness	No evidence	
Pain		
Comparative efficacy	Low	Evidence from one fair placebo-controlled study is insufficient to draw conclusions about treating depression in patients with coexisting pain. Results from head-to-head trials are not available.
Comparative effectiveness	No evidence	
Psychomotor change		
Comparative efficacy	Low	Evidence from one fair head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting psychomotor change. Results indicate comparative outcomes for psychomotor retardation and psychomotor change may be different.
Comparative effectiveness	No evidence	
Somatization		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Key Question 3b: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of symptom clusters in patients with depression		
Anxiety		
Comparative efficacy	Moderate	Results from 10 fair head-to-head trials and 2 fair placebo-controlled trials suggest that no substantial differences in efficacy exist among second-generation antidepressants for treatment of accompanying anxiety symptoms.
Comparative effectiveness	No evidence	
Insomnia		
Comparative efficacy	Low	Evidence from six fair head-to-head trials is insufficient to draw conclusions about treating insomnia in depressed patients. Results are limited by study design; differences in outcomes are of unknown clinical significance.
Comparative effectiveness	No evidence	
Melancholia		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Pain		
Comparative efficacy	Low	Evidence from two head-to-head trials (one fair, one poor) and three placebo-controlled trials is insufficient to draw conclusions about treating coexisting pain in depressed patients. Results indicate no difference in efficacy but are limited by study design.
Comparative effectiveness	No evidence	

Table 32. Summary of findings with strength of evidence (continued)

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Psychomotor change		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Somatization		
Comparative efficacy	No evidence	
Comparative effectiveness	Low	Evidence from one open-label head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating coexisting somatization in depressed patients. Results indicate no difference in effectiveness.
Key Question 4: Comparative risk of harms (safety, adverse events) and adherence		
General tolerability		
Adverse events profiles	High	Adverse events profiles are similar among second-generation antidepressants. Differences in the incidence of specific adverse events exist.
Nausea and vomiting	High	Meta-analysis of 15 fair-quality studies indicates that venlafaxine has a higher rate of nausea and vomiting than SSRIs as a class.
Diarrhea	Moderate	Evidence from 15 fair-quality studies indicates that sertraline has a higher incidence of diarrhea than bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, and venlafaxine.
Weight change	Moderate	Seven fair trials indicate that mirtazapine leads to higher weight gains than citalopram, fluoxetine, paroxetine, and sertraline.
Somnolence	Moderate	Six fair studies provide evidence that trazodone has a higher rate of somnolence than bupropion, fluoxetine, mirtazapine, paroxetine, and venlafaxine.
Discontinuation syndrome	Moderate	A good systematic review provides evidence that paroxetine and venlafaxine have the highest rates of discontinuation syndrome; fluoxetine has the lowest.
Discontinuation rates	High	Meta-analyses of efficacy trials indicate that overall discontinuation rates are similar. Venlafaxine has a higher rate of discontinuations because of adverse events and a lower rate of discontinuations because of lack of efficacy than SSRIs as a class.
Severe adverse events		
Suicidality	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of suicidality.
Sexual adverse events	Moderate	Five fair trials provide evidence that bupropion causes significantly less sexual dysfunction than fluoxetine, paroxetine, and sertraline. Among SSRIs, paroxetine has the highest rates of sexual dysfunction.
Cardiovascular adverse events	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of cardiovascular adverse events. Weak evidence indicates that venlafaxine might have an increased risk of cardiovascular adverse events.
Hyponatremia	Low	The evidence is insufficient to draw conclusions about the comparative risk for hyponatremia.

Table 32. Summary of findings with strength of evidence (continued)

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Seizures	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of seizures. Weak evidence indicates that bupropion might have an increased risk of seizures.
Hepatotoxicity	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of hepatotoxicity. Weak evidence indicates that nefazodone might have an increased risk of hepatotoxicity.
Serotonin syndrome	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of serotonin syndrome. Observational studies indicate no differences in risk among second-generation antidepressants.
Adherence in efficacy studies	Moderate	Efficacy studies indicate no differences in adherence. One observational study suggests that extended-release formulations might have better adherence than immediate-release formulations.
Adherence in effectiveness studies	Low	Evidence from existing studies is insufficient to draw conclusions about adherence in "real-world" settings.
Key Question 5: Subgroups		
Age		
Comparative efficacy	Moderate	Results from 22 efficacy trials (2 good RCTs, 17 fair RCTs or pooled analyses of RCTs, 1 poor RCT, 1 pooled analysis that was not rated, and 1 nonrandomized controlled trial that was not rated) indicate that no substantial differences exist in efficacy among second-generation antidepressants in the elderly or the very elderly.
Comparative effectiveness	Moderate	Based on findings from one fair head-to-head effectiveness trial, no substantial differences exist among second-generation antidepressants in the elderly compared with other age groups. A second trial in patients with dysthymia or minor depression provides mixed evidence.
Comparative harms	Low	Results from two fair studies indicate that adverse events may differ somewhat across second-generation antidepressants in the elderly or very elderly.
Sex		
Comparative efficacy	Low	Results from one fair pooled analysis of RCTs indicates that efficacy among second-generation antidepressants may not differ substantially between men and women.
Comparative effectiveness	No evidence	
Comparative harms	Low	One fair head-to-head trial suggests harms (headache, nausea) may differ between men and women treated with venlafaxine vs. placebo and venlafaxine vs. SSRIs or placebo. Observational evidence (one fair study) suggests that some sexual side effects may differ between in men and women.

Table 32. Summary of findings with strength of evidence (continued)

Key Question, Disorder, and Outcome of Interest	Strength of Evidence	Findings
Race or Ethnicity		
Comparative efficacy	Low	Results from one poor RCT indicate that efficacy does not differ substantially among second-generation antidepressants in different racial subgroups.
Comparative effectiveness	No evidence	
Comparative harms	No evidence	
Comorbidities		
Comparative efficacy	Low	One poor head-to-head trial included patients with depression and HIV/AIDS; this study indicated that efficacy does not differ substantially among second-generation antidepressants. Findings from placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	No evidence	
Comparative harms	No evidence	

Although second-generation antidepressants are similar in efficacy, they cannot be considered identical drugs. Evidence of moderate strength supports some differences among individual drugs with respect to onset of action, adverse events, and some measures of health-related quality of life; these are of modest magnitude but statistically significant. Specifically, consistent evidence from multiple trials demonstrates that mirtazapine has a faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline[60,61,72,73,76] and that bupropion has fewer sexual side effects than fluoxetine, paroxetine, and sertraline.[79,80,88-90]

Some of these differences are small and might be offset by other adverse events. For example, a faster onset of mirtazapine must be weighed against possible decreased adherence because of long-term weight gain. Nonetheless, some of these differences may be clinically significant and influence the choice of a medication for specific patients. For example, patients who have a history of nausea or who dread sexual dysfunction might be more adherent to a choice of treatment that takes these factors into consideration. Past treatment experiences may also frame decisions regarding medications to either select or avoid, but no evidence exists to verify these inferences.

A considerable limitation of our conclusions is that they have been derived primarily from efficacy trials. Although findings from effectiveness studies are generally consistent with those from efficacy trials, the generalizability of some of our conclusions may be limited. Furthermore, the pharmaceutical industry funded a large percentage of these studies, and selective reporting is conceivable, although we had no way to account for missing information.

Our report is the first to assess statistically each of 66 possible drug comparisons of second-generation antidepressants. For comparative efficacy, we employed direct analyses for four comparisons and 62 indirect statistical analyses.

In the following sections we discuss major findings for individual key questions in more detail.

Results for Efficacy and Effectiveness in Major Depressive Disorders

For MDD, direct evidence from head-to-head trials and indirect comparisons using placebo-controlled trials indicate that, overall, the efficacy and effectiveness of second-generation antidepressants do not differ substantially for the treatment of adults. We rated the strength of this evidence as moderate. These findings are consistent with prior systematic reviews and meta-analyses.[8,241]

In some of our meta-analyses, results of pooled response rates indicate statistically significant differences in efficacy between some drugs. Specifically, for response, escitalopram is more efficacious than citalopram, sertraline more than fluoxetine, and venlafaxine more than fluoxetine. Accompanying meta-analyses of effect sizes, however, suggest that the actual differences in the mean treatment effects are small and most likely not clinically significant.

For example, a relative risk (RR) meta-analysis of response rates indicates that significantly more patients receiving escitalopram than receiving citalopram achieved treatment response (RR, 1.14; 95% CI, 1.04-1.26). An effect-size meta-analysis yielded a mean difference of 1.3 points on the Hamilton Depression Rating Scale (HAM-D), which represents about one-fifth to one-quarter of a standard deviation. Therefore, this difference most likely does not represent a minimal clinically significant difference. A recent methods study concluded that a change of about one-half of a standard deviation reflects a minimal important difference for a patient.[103] In this case, dichotomizing a continuous scale such as the HAM-D appears to overestimate the actual difference in effect sizes.

Similarly, sertraline and venlafaxine had statistically significantly greater response rates than fluoxetine. Effect size meta-analyses, however, yielded no clinically significant mean differences on HAM-D scales.

Findings from indirect comparisons yielded no statistically significant differences in response rates among other potential comparisons. The precision of some of these estimates was low, leading to inconclusive results with wide confidence intervals. Nevertheless, point estimates of treatment effects consistently indicate no substantial differences in efficacy among comparisons.

Although response and remission rates are similar among second-generation antidepressants, 54 percent of patients in these trials did not achieve remission and 34 percent did not respond. Many of these patients will require a second-line treatment. Results from the Sequenced Treatment Alternatives to Relieve Depression (STAR-D) trial—an effectiveness study that randomized patients to bupropion SR, sertraline, or venlafaxine XR after they had failed treatment with citalopram[146]—indicate that, even with second-line treatments, a substantial proportion of patients do not achieve remission.

Effectiveness trials have greater generalizability of findings than efficacy studies; we found only three such trials. Two of these effectiveness trials were conducted in French primary care settings and one was performed in the United States. Findings were generally consistent with efficacy trials—they did not detect any substantial differences in effectiveness. However, differences between French and US health systems may limit the applicability of results from French effectiveness trials to US patients.

No evidence exists on adherence in effectiveness studies. Although adherence was similar in efficacy trials, the generalizability of such findings may be limited. Most likely, dosing regimens, adverse events, and costs substantially influence adherence of patients in everyday

practice. Given similar efficacy and effectiveness, such factors need to be considered when choosing a medication.

Results for Maintaining Response or Remission

The majority of studies included in this report involved treating patients with major depression in its acute phase; for this phase, the goal is reducing signs and symptoms of depression to achieve remission. Patients who achieve remission with acute-phase treatment should be followed to maintain that response and remission. That is, they should be managed in a continuation phase to prevent relapse and, if necessary, in a longer-term maintenance phase to prevent recurrence. (See Figure 1 in the introduction for clarification of these treatment cycles.)

Although evidence was sparse on the comparative efficacy and effectiveness for maintaining response or remission, treating recurrent depression, or treating depression that does not respond to first-line treatment, our findings are consistent with results from acute-phase trials. Overall, no substantial differences among second-generation antidepressants were apparent, but comparisons are limited to a few drugs.

Moderate strength evidence from three efficacy trials[47,96,116,117] suggests that no substantial differences in efficacy exist between fluoxetine and sertraline, fluvoxamine and sertraline, and trazodone and venlafaxine for preventing relapse or recurrence. Although results are consistent across these studies, evidence for other drug comparisons is not available; hence, these results are not generalizable to other second-generation antidepressants.

Additionally, trials differed in their design and conduct, further limiting the applicability (generalizability) of this evidence. For example, criteria used to define relapse and recurrence differed considerably across trials. As cases in point with respect to relapse: In the three head-to-head studies, one defined relapse as an increase in the lowest HAM-D or Montgomery-Asberg Depression Rating Scale (MADRS) score of at least 50 percent for 2 weeks, a HAM-D greater than 18 for 2 weeks, and a Clinical Global Impressions – Severity (CGI-S) score greater than 4;[47] a second study defined relapse as a HAM-D score greater than 15 with functional impairment;[116,117] and the third simply assessed discontinuation rates.[96] Eligibility for continuation- or maintenance-phase treatment also varied considerably.

We advise that, in future studies, investigators try to build on past and current work by employing definitions of relapse that are similar to those commonly found in the published literature to date. In our view, convergence on standard, accepted definitions of recurrence would be useful as well.

A related question may be how long to continue treatment intended to prevent relapse and recurrence. Although we did not set out to answer this question, we believe that some evidence suggests that the risk of relapse decreases over time. For example, one placebo-controlled study compared 14 weeks, 38 weeks, and 50 weeks of continuation treatment with fluoxetine or placebo.[122] Relapse rates were significantly lower for patients on fluoxetine than for those on placebo at 14 and 38 weeks, but not at 50 weeks. This finding implies some degree of diminishing returns for longer treatment, although more work is needed to address this question.

Results for Managing Treatment-Resistant or Recurrent Depression

Overall, approximately 40 percent of patients do not achieve clinical response with initial treatment; approximately 10 percent to 15 percent of patients discontinue treatment because of

adverse events. Three studies addressed the comparative efficacy or effectiveness among second-generation antidepressants in patients with treatment-resistant depression. These studies came to inconsistent conclusions, although some of these inconsistencies may be partially explained by variations in the quality and applicability (i.e., internal and external validity) of these investigations. We rated the strength of evidence as moderate.

The best evidence comes from the STAR-D trial.[146] Although this was an open-label study, an interviewer blinded to the treatment arm did the outcomes assessment. Among patients who did not have a remission or could not tolerate citalopram, the investigators reported that bupropion SR, sertraline, and venlafaxine XR had similar effectiveness and tolerability as second-line treatment. Although the ARGOS study, another effectiveness study, found venlafaxine to be superior to citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline as a second-step treatment,[140] we could not determine whether raters were blinded to treatment allocation, potentially limiting the ARGOS conclusions.

No study specifically compared one antidepressant with another in patients experiencing a depressive relapse (i.e., loss of response during continuation-phase treatment) or recurrence (i.e., loss of response during maintenance-phase treatment). Although STAR-D included patients with a history of recurrent depressive episodes at study entry, the analyses involved patients whose acute-phase treatment of the current episode had been unsuccessful; it did not include patients who initially responded and then lost response.

Results for Treating Patients with Depression and Accompanying Symptoms

The range of physical and psychological symptoms that accompany depressive disorders is wide. We found limited information for many accompanying symptom clusters; however, various symptoms may not have the same importance for clinical care. Our analyses concerned the efficacy and effectiveness of these pharmaceuticals for treating depression in patients with such symptoms *and* treating the accompanying symptoms in patients with depression. Generally, the strength of evidence for anxiety was moderate; for all other symptom clusters, either the strength of evidence was low or no evidence was found.

The most common and distressing accompanying symptoms can be considered the highest priority for further studies. Research involving depressed populations that may be more generalizable suggests that common presenting symptom clusters in both primary care and psychiatric clinics are fatigue and loss of energy (for which no studies were identified), anxiety, insomnia, and pain and other somatic symptoms.[147]

Anxiety. Although anxiety is not a discrete MDD subtype,[242] evidence suggests that it may present as a distinctive cluster[243] and be associated with more persistent depression.[244-246] For patients with high anxiety associated with MDD, we found no difference in patients' depression treatment response by either antidepressant class or specific medication. These findings are consistent with a recent nonsystematic review sponsored by a pharmaceutical manufacturer.[247] Although all the included studies identified a high anxiety group, the definitions employed by investigators varied markedly.

In addition, for patients with anxiety symptoms associated with depression, we found no identifiable difference in anxiety response by either antidepressant class or specific medication. Therefore, the current evidence suggests that improvement in both depressive and anxiety

symptoms is likely with adequate dosing of antidepressant treatment, but evidence of clear benefit for one antidepressant over another is lacking.

Insomnia. For patients with depression and accompanying insomnia, we found no clear evidence of differences in depressive response or insomnia response by antidepressant class or specific medication.

Indirect evidence from studies that did not identify insomnia subgroups[82,96] provides results that are consistent with improved sleep quality for trazodone compared with fluoxetine[82] and venlafaxine.[96] Higher quality, direct evidence, however, was limited. Among the three studies that identified an insomnia group, only one trial involved one of these three antidepressants; it suggested greater benefit for nefazodone than fluoxetine.[81] The two other studies, which compared SSRIs, produced mixed results.[41,151]

Studies were limited by varying and incomplete assessment of insomnia and by insensitive outcome measures. Most studies used a sleep measure that is a part of HAM-D, with three items producing a total sleep score ranging from 0 to 6. The clinical meaningfulness of the small reported differences in this outcome measure is unclear.

Melancholia. Information about outcomes in the melancholic subgroup was limited to three comparative trials; they addressed only the effect on depressive outcomes. Evidence did not consistently support a difference in outcome by either class or medication.

Pain. Patients with depression commonly experience physical symptoms; the majority are pain symptoms. In addition, depression is prevalent among patients with chronic pain disorders.[248] We identified few trials addressing the use of second-generation antidepressants for treatment of pain accompanying depression. All the trials we identified tested duloxetine, an SSNRI; two compared duloxetine with paroxetine, and the other three were placebo-controlled trials.

Studies were limited by exclusion of patients with common chronic pain conditions, failure to analyze subgroups with moderate to severe pain, and failure to report outcomes in a clinically meaningful way. No study included patients with comorbid depression and chronic pain, probably the group of most interest to clinicians. The only study that required patients to have pain of at least mild intensity for inclusion excluded those with a history of any diagnosed painful condition, including common pain disorders such as migraine and arthritis.[154]

The difference in mean pain scores between duloxetine and placebo groups was statistically significant, but probably not clinically meaningful, in three studies; all used a 100 mm pain intensity visual analog scale (VAS) as the outcome measure.[153,155,156] Prior research has produced different estimates of the minimum clinically important difference on the VAS, ranging from 9 mm to 30 mm.[109,249-251] No study included in this review reported the proportion of patients achieving a clinically important improvement in pain scores.

Psychomotor changes. The evidence addressing depression outcomes in patients with psychomotor changes is limited to a single trial. It found that sertraline was more efficacious than fluoxetine in patients with psychomotor agitation but not in those with psychomotor retardation.[149]

Somatization. The evidence directly addressing treatment of somatization in patients with depression is limited to a single trial that found similar effectiveness for three SSRIs.[49] Conclusions from this study are limited because the investigators did not analyze information for a subgroup with high somatization.

Results for Harms (Adverse Events) and Adherence

On average, 61 percent of patients experienced at least one adverse event during the course of the studies we reviewed. Nausea, headache, diarrhea, fatigue, dizziness, sweating, tremor, dry mouth, and weight gain were commonly reported adverse events.

Although the spectrum of adverse events is similar among second-generation antidepressants, the frequencies of specific adverse events differ among individual drugs. For example, venlafaxine had a higher rate of nausea and vomiting than the SSRIs as a class. Also, compared with other second-generation antidepressants, paroxetine frequently led to higher sexual side effects, mirtazapine and paroxetine to higher weight gains, and sertraline to a higher rate of diarrhea. Such differences did not lead to substantial differences in discontinuation rates.

For some patients, these differences might well be clinically important. For example, the choice of an agent with a low rate of sexual side effects might increase adherence in patients who consider sexual dysfunction an intolerable adverse event.

The evidence on the comparative risk for rare but severe adverse events such as suicidality, hyponatremia, seizures, or serotonin syndrome was insufficient to draw firm conclusions. The risk of such harms should be kept in mind during any course of treatment with a second-generation antidepressant.

Efficacy studies did not indicate any differences in adherence across agents. One observational study indicated that extended-release formulations might have a better adherence rate than immediate-release medications. This finding, however, is likely more attributable to differences in dosing regimens than to differences in efficacy and harms. The evidence is insufficient to draw any conclusions about differences in adherence in effectiveness studies.

Results for Population Subgroups

In efficacy and effectiveness studies, treatment effects were similar between different age groups and between males and females. Despite the importance of the harms of second-generation antidepressants, especially in the elderly, little evidence is available on this topic. We found very limited head-to-head evidence assessing potential differences in efficacy in different racial groups or in patients with common comorbidities. Specifically for different racial groups and for patients with common comorbidities, the evidence is sparse and mainly limited to placebo-controlled trials assessing the general efficacy of second-generation antidepressants in such subgroups. Some of these studies indicate that the general efficacy of second-generation antidepressants in patients with serious comorbidities (e.g., cancer, substance abuse) is limited.

Many of these studies had serious methodological flaws or were too small to detect meaningful differences, although they may not have been powered to detect significant differences. Differences in study populations, cutoff points on scales, and drug dosages do not allow analysts to compare initial treatment effects across individual placebo-controlled trials to assess differences in subgroups other than those defined by age and sex.

Results for Dysthymia and Subsyndromal Depression

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

For the treatment of dysthymia, the evidence on general efficacy is limited to fluoxetine, paroxetine, and sertraline; for subsyndromal depression, the evidence covers only citalopram, fluoxetine, and paroxetine. Results are mixed. For dysthymia, the two largest placebo-controlled studies did not detect any differences between fluoxetine or paroxetine and placebo for treating patients younger than 60 years.[100,113] Similarly, the evidence on the general efficacy in subsyndromal depression is limited to few studies with mixed results.

Future Research

We identified multiple areas that require additional research to enable clinicians and researchers to draw firm conclusions about the comparative efficacy, effectiveness, and harms of second-generation antidepressants.

Efficacy and Effectiveness

Future research has to establish reliably the general efficacy of second-generation antidepressants for the treatment of dysthymia and subsyndromal depression. Ideally, multiple-arm, head-to-head trials, including placebo groups, should evaluate the general and comparative efficacy of second-generation antidepressants in patients with these conditions.

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

Future research should also focus on differences in efficacy and effectiveness in subgroups such as the very elderly or patients with various common comorbidities.

Prevention of Relapse and Recurrence

More evidence is needed regarding the most appropriate duration of antidepressant treatment for maintaining remission. Such studies should also evaluate whether different formulations (i.e., controlled release vs. immediate release) lead to differences in adherence and subsequently to differences in relapse or recurrence.

Additionally, although most trials maintained the dose used in acute-phase treatment throughout continuation and maintenance treatment, little is known about the effect of drug dose on the risk of relapse or recurrence. The effect of differences in drug doses is also poorly understood.

Management of Treatment-Resistant or Recurrent Depression

Given the fact that approximately 40 percent of patients do not respond to initial treatment, an important future research agenda is to explore whether combinations of antidepressants at treatment initiation lead to better response rates than single agents alone. Furthermore, additional head-to-head evidence is needed to resolve whether one second-generation antidepressant is better than another in patients who either did not respond or could not tolerate a first-line treatment.

Likewise, evidence is lacking to determine whether one antidepressant is better than another in patients who cannot maintain remission during continuation- or maintenance-phase therapy. The role of other depression treatments, such as psychotherapy, vagal nerve stimulation, light therapy, and alternative medicines as substitutes or complements to pharmaceutical management also needs to be better understood.

Accompanying Symptoms

More research is needed to evaluate differences between second-generation antidepressants in populations with accompanying symptoms such as anxiety, insomnia, pain, and fatigue. Given that outcomes for depression treatment do not differ substantially between specific antidepressants, information about treatment of accompanying symptoms is key for clinicians who must select among many antidepressant drugs.

Study questions must be based on a clinically meaningful metric that gives preference to symptoms of high frequency or those that cause a high level of distress. Each subgroup must be clearly and consistently defined (e.g., a high anxiety group should be identified with a consistent definition). Analyses should then be done in such subgroups, using similarly defined outcomes to allow results to be compared across studies and across subgroups. Investigators should report the proportions of patients who reach a predefined threshold for clinically meaningful improvement.

The absence of any trials conducted in a population with fatigue or loss of energy presents a clinically important void in the literature. In addition, future studies of depression with accompanying pain and other somatic symptoms should identify clinically relevant subgroups of patients with moderate to severe pain or other symptoms.

Adverse Events

Large, well-conducted observational studies are needed to assess reliably the comparative risks of second-generation antidepressants with respect to rare but serious adverse events such as suicidality, hyponatremia, hepatotoxicity, seizures, cardiovascular adverse events, and serotonin syndrome. Furthermore, these studies need to evaluate whether very elderly patients have an excess risk of severe adverse events with any second-generation antidepressant.

Addendum

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

The Sequenced Treatment Alternatives to Relieve Depression (STAR-D) trial – described in detail in Key Question 2b – consisted of a series of RCTs examining sequential treatment steps in patients who did not obtain remission or could not tolerate previous treatments. Key Question 2b detailed the medication switch arms of the second-step treatment in which all patients in the analysis had failed initial treatment with citalopram and were randomized to second-step treatment with bupropion SR (N = 239), sertraline (N = 238), or venlafaxine XR (N = 250); this analysis found no statistically significant differences in remission rates between second-step treatments.[146]

The more recently published study describes the acute and longer-term outcomes associated with all four treatment steps.[252] Patients not achieving remission or unable to tolerate a treatment step were encouraged to move to the next step; patients achieving acceptable benefit could enter a 12-month follow-up phase. All patients (N = 3,671) received citalopram in Step 1. Step 2 and Step 3 treatments were randomly assigned using an equipoise stratified randomized design. In this, 1,439 patients were randomized in Step 2, which included seven possible treatment alternatives (bupropion SR, sertraline, venlafaxine XR, cognitive therapy, citalopram plus bupropion, citalopram plus buspirone, or citalopram plus cognitive therapy). Step 3 randomized 390 patients to switch to mirtazapine or nortriptyline or to receive augmentation with lithium or triiodothyronine (T3). Step 4 used only a single randomization; 123 patients were randomized to tranylcypromine or venlafaxine XR plus mirtazapine.

Overall, 67 percent of patients achieved remission. Remission rates were 36.8 percent for Step 1, 30.6 percent for Step 2, 13.7 percent for Step 3, and 13.0 percent for Step 4. For patients achieving acceptable benefits who continued on in the 12-month follow-up study, relapse rates were 40.1 percent, 55.3 percent, 64.6 percent, and 71.1 percent for those achieving benefit in Steps 1, 2, 3, and 4, respectively. In all steps, patients achieving remission (Quick Inventory of Depressive Symptomatology–Self Report [QIDS-SR-16] $\leq$ 5) were less likely to relapse than patients not achieving remission (acceptable benefit but QIDS-SR-16 > 5).

References

1. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry. 1994;51(1):8-19.

2. 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). JAMA. 2003;289(23):3095-105.

3. 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(12):1465-75.

4. US Food and Drug Administration. Electronic Orange Book. 2004 [cited; Available from: http://www.fda.gov/cder/ob/default.htm

5. Maj J, Palider W, Rawlow. Trazodone, a central serotonin antagonist and agonist. J Neural Transm. 1979;44(3):237-48.

6. Stefanini E, Fadda F, Medda L, Gessa GL. Selective inhibition of serotonin uptake by trazodone, a new antidepressant agent. Life Sci. 1976 Jun 15;18(12):1459-65.

7. IMS Health. Press Release: Growth is sustained by new products despite a difficult year http://www.imshealth.com/ims/portal/front/articleC/0,2777,6599_3665_44771558,00.html. IMS Reports; 2004.

8. Williams JW, Mulrow CD, Chiquette E, Noel PH, Augilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. Ann Intern Med. 2000;132(9):743-56.

9. Geddes JR, Freemantle N, Mason J, Eccles MP, Boynton J. Selective serotonin reuptake inhibitors (SSRIs) versus other antidepressants for depression. The Cochrane Library (Cochrane Review). 2006(1).

10. Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E, et al. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. Lancet. 2003 Feb 22;361(9358):653-61.

11. Kupfer DJ, Perel JM, Pollock BG, Nathan RS, Grochocinski VJ, Wilson MJ, et al. Fluvoxamine versus desipramine: comparative polysomnographic effects. Biol Psychiatry. 1991 Jan 1;29(1):23-40.

12. American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder (revision). Am J Psychiatry. 2000 Apr;157(4 Suppl):1-45.

13. Depression Guideline Panel. Depression in primary care: Volume 2, Treatment of major depression. . Rockville, MD: Vol AHCPR publication No. 93-0550: US DHHS, Public Health Service, Agency for Health Care Policy and Research; 1993.

14. Hoagwood K, Hibbs E, Brent D, Jensen P. Introduction to the special section: efficacy and effectiveness in studies of child and adolescent psychotherapy. J Consult Clin Psychol. 1995;63(5):683-7.

15. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999;354(9193):1896-900.

16. 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. Am J Prev Med. 2001;20(3 Suppl):21-35.

17. Anonymous. Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. CRD Report Number 4 (2nd edition). 2001.

18. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. Health Technol Assess. 2003;7(27):iii-x, 1-173.

19. Gartlehner G, Hansen RA, Nissman D, Lohr KN, Carey TS. A simple and valid tool distinguished efficacy from effectiveness studies. J of Clinical Epidemiology. 2006;59(10):1040-8.

20. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. BMC Health Serv Res. 2004 Dec 22;4(1):38.

21. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, et al. Grading strength of recommendations and quality of evidence

in clinical guidelines: report from an American College of Chest Physicians task force. Chest. 2006 Jan;129(1):174-81.

22. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. Health Technol Assess. 2005 Jul;9(26):1-134, iii-iv.

23. Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med. 2002 Aug 30;21(16):2313-24.

24. Burke WJ, Gergel I, Bose A. Fixed-dose trial of the single isomer SSRI escitalopram in depressed outpatients. J Clin Psychiatry. 2002 Apr;63(4):331-6.

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

26. 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 Jul;18(4):211-7.

27. Moore N, Verdoux H, Fantino B. Prospective, multicentre, randomized, double-blind study of the efficacy of escitalopram versus citalopram in outpatient treatment of major depressive disorder. Int Clin Psychopharmacol. 2005 May;20(3):131-7.

28. 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 Jun;11(2):129-36.

29. Haffmans PM, Timmerman L, Hoogduin CA. Efficacy and tolerability of citalopram in comparison with fluvoxamine in depressed outpatients: a double-blind, multicentre study. The LUCIFER Group. Int Clin Psychopharmacol. 1996 Sep;11(3):157-64.

30. 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 Nov;12(6):323-31.

31. Kasper S, de Swart H, Friis Andersen H. Escitalopram in the treatment of depressed elderly patients. Am J Geriatr Psychiatry. 2005 Oct;13(10):884-91.

32. Dalery J, Honig A. Fluvoxamine versus fluoxetine in major depressive episode: a double-blind randomised comparison. Hum Psychopharmacol. 2003 Jul;18(5):379-84.

33. 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 Oct;16(5):373-8.

34. Cassano GB, Puca F, Scapicchio PL, Trabucchi M. Paroxetine and fluoxetine effects on mood and cognitive functions in depressed nondemented elderly patients. J Clin Psychiatry. 2002 May;63(5):396-402.

35. 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 Jul;54(1-2):39-48.

36. 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 Feb;87(2):141-5.

37. 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 Dec;10(4):145-50.

38. Gagiano CA. A double blind comparison of paroxetine and fluoxetine in patients with major depression. Br J Clin Res. 1993;4:145-52.

39. Schone W, Ludwig M. A double-blind study of paroxetine compared with fluoxetine in geriatric patients with major depression. J Clin Psychopharmacol. 1993 Dec;13(6 Suppl 2):34S-9S.

40. Tignol J. A double-blind, randomized, fluoxetine-controlled, multicenter study of paroxetine in the treatment of depression. J Clin Psychopharmacol. 1993 Dec;13(6 Suppl 2):18S-22S.

41. 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 Apr;22(2):137-47.

42. Bennie EH, Mullin JM, Martindale JJ. A double-blind multicenter trial comparing sertraline and fluoxetine in outpatients with major depression. J Clin Psychiatry. 1995 Jun;56(6):229-37.

43. Boyer P, Danion JM, Bisserbe 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 Jan;13(1 Pt 2):157-69.

44. Newhouse PA, Krishnan KR, Doraiswamy PM, Richter EM, Batzar ED, Clary CM. A double-blind comparison of sertraline and fluoxetine in depressed

elderly outpatients. J Clin Psychiatry. 2000 Aug;61(8):559-68.

45. Finkel SI, Richter EM, Clary CM, Batzar E. Comparative efficacy of sertraline vs. fluoxetine in patients age 70 or over with major depression. Am J Geriatr Psychiatry. 1999 Summer;7(3):221-7.

46. 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 Mar;14(1):41-8.

47. Van Moffaert M, Bartholome F, Cosyns P, De Nayer AR. A controlled comparison of sertraline and fluoxetine in acute and continuation treatment of major depression. Hum Psychopharmacol. 1995;10:393-405.

48. 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 Nov;61(11):863-7.

49. Kroenke K, West SL, Swindle R, Gilsenan A, Eckert GJ, Dolor R, et al. Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial. JAMA. 2001 Dec 19;286(23):2947-55.

50. Kiev A, Feiger A. A double-blind comparison of fluvoxamine and paroxetine in the treatment of depressed outpatients. J Clin Psychiatry. 1997 Apr;58(4):146-52.

51. 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(4):163-9.

52. Rossini D, Serretti A, Franchini L, Mandelli L, Smeraldi E, De Ronchi D, et al. Sertraline versus fluvoxamine in the treatment of elderly patients with major depression: a double-blind, randomized trial. J Clin Psychopharmacol. 2005 Oct;25(5):471-5.

53. 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 Dec;20(6):645-52.

54. Leinonen E, Skarstein J, Behnke K, Agren H, Helsdingen JT. Efficacy and tolerability of mirtazapine versus citalopram: a double-blind, randomized study in patients with major depressive disorder. Nordic Antidepressant Study Group. Int Clin Psychopharmacol. 1999 Nov;14(6):329-37.

55. Allard P, Gram L, Timdahl K, Behnke K, Hanson M, Sogaard J. Efficacy and tolerability of venlafaxine in geriatric outpatients with major depression: a double-blind, randomised 6-month comparative trial with citalopram. Int J Geriatr Psychiatry. 2004 Dec;19(12):1123-30.

56. Bielski RJ, Ventura D, Chang CC. A double-blind comparison of escitalopram and venlafaxine extended release in the treatment of major depressive disorder. J Clin Psychiatry. 2004 Sep;65(9):1190-6.

57. 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(1):57-64.

58. 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 Mar;63(3):225-31.

59. 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 Aug;64(8):921-6.

60. Versiani M, Moreno R, Ramakers-van Moorsel CJ, Schutte AJ. Comparison of the effects of mirtazapine and fluoxetine in severely depressed patients. CNS Drugs. 2005;19(2):137-46.

61. Wheatley DP, van Moffaert M, Timmerman L, Kremer CM. Mirtazapine: efficacy and tolerability in comparison with fluoxetine in patients with moderate to severe major depressive disorder. Mirtazapine-Fluoxetine Study Group. J Clin Psychiatry. 1998 Jun;59(6):306-12.

62. Alves C, Cachola I, Brandao J. Efficacy and tolerability of venlafaxine and fluoxetine in outpatients with major depression. Primary Care Psychiatry. 1999;5(2):57-63.

63. Costa e Silva J. Randomized, double-blind comparison of venlafaxine and fluoxetine in outpatients with major depression. J Clin Psychiatry. 1998 Jul;59(7):352-7.

64. 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 Jun;5(2):115-20.

65. 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 Jan;20(1):57-71.

66. Nemeroff CB, Thase ME. A double-blind, placebo-controlled comparison of venlafaxine and fluoxetine treatment in depressed outpatients. J Psychiatr Res. 2005 Sep 12.

67. 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 Dec;56(2-3):171-81.

68. 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 Jan;60(1):22-8.

69. Tzanakaki M, Guazzelli M, Nimatoudis I, Zissis NP, Smeraldi E, Rizzo F. Increased remission rates with venlafaxine compared with fluoxetine in hospitalized patients with major depression and melancholia. Int Clin Psychopharmacol. 2000 Jan;15(1):29-34.

70. 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 major depression in general practice. Primary Care Psychiatry. 1997;3(1):51-8.

71. 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 Dec;14(6):457-70.

72. Benkert O, Szegedi A, Kohnen R. Mirtazapine compared with paroxetine in major depression. J Clin Psychiatry. 2000 Sep;61(9):656-63.

73. Schatzberg AF, Kremer C, Rodrigues HE, Murphy GM, Jr. Double-blind, randomized comparison of mirtazapine and paroxetine in elderly depressed patients. Am J Geriatr Psychiatry. 2002 Sep-Oct;10(5):541-50.

74. 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 Jan;15(1):43-8.

75. 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 1998;4(3):127-32.

76. 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 Aug;23(4):358-64.

77. 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 Feb;61(2):95-100.

78. Sir A, D'Souza RF, Uguz S, George T, Vahip S, Hopwood M, et al. Randomized trial of sertraline versus venlafaxine XR in major depression: efficacy and discontinuation symptoms. J Clin Psychiatry. 2005 Oct;66(10):1312-20.

79. 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 Jul;23(7):1040-58.

80. 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 Aug;52(8):329-35.

81. Rush AJ, Armitage R, Gillin JC, Yonkers KA, Winokur A, Moldofsky H, et al. Comparative effects of nefazodone and fluoxetine on sleep in outpatients with major depressive disorder. Biol Psychiatry. 1998 Jul 1;44(1):3-14.

82. Beasley CM, Jr., Dornseif BE, Pultz JA, Bosomworth JC, Sayler ME. Fluoxetine versus trazodone: efficacy and activating-sedating effects. J Clin Psychiatry. 1991 Jul;52(7):294-9.

83. Perry PJ, Garvey MJ, Kelly MW, Cook BL, Dunner FJ, Winokur G. A comparative trial of fluoxetine versus trazodone in outpatients with major depression. J Clin Psychiatry. 1989 Aug;50(8):290-4.

84. 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 Mar;61(3):196-202.

85. Baldwin DS, Hawley CJ, Abed RT, Maragakis BP, Cox J, Buckingham SA, et al. A multicenter double-blind comparison of nefazodone and paroxetine in the treatment of outpatients with moderate-to-severe depression. J Clin Psychiatry. 1996;57 Suppl 2:46-52.

86. Hicks JA, Argyropoulos SV, Rich AS, Nash JR, Bell CJ, Edwards C, et al. Randomised controlled study of sleep after nefazodone or paroxetine treatment in out-patients with depression. Br J Psychiatry. 2002 Jun;180:528-35.

87. Kasper S, Olivieri L, Di Loreto G, Dionisio P. A comparative, randomised, double-blind study of

trazodone prolonged-release and paroxetine in the treatment of patients with major depressive disorder. Curr Med Res Opin. 2005 Aug;21(8):1139-46.

88. Coleman CC, Cunningham LA, 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 Dec;11(4):205-15.

89. 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 Apr;21(4):643-58.

90. 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 Dec;58(12):532-7.

91. Rush AJ, Batey SR, Donahue RM, Ascher JA, Carmody TJ, Metz A. Does pretreatment anxiety predict response to either bupropion SR or sertraline? J Affect Disord. 2001 Apr;64(1):81-7.

92. Feiger A, Kiev A, Shrivastava RK, Wisselink PG, Wilcox CS. Nefazodone versus sertraline in outpatients with major depression: focus on efficacy, tolerability, and effects on sexual function and satisfaction. J Clin Psychiatry. 1996;57 Suppl 2:53-62.

93. Guelfi JD, Ansseau M, Timmerman L, Korsgaard S. Mirtazapine versus venlafaxine in hospitalized severely depressed patients with melancholic features. J Clin Psychopharmacol. 2001 Aug;21(4):425-31.

94. Halikas JA. Org 3770 (mirtazapine) versus trazodone: A placebo controlled trial in depressed elderly patients. Hum Psychopharmacol. 1995;10(Suppl 2):S125-S33.

95. van Moffaert M, de Wilde J, Vereecken A, Dierick M, Evrard JL, Wilmotte J, et al. Mirtazapine is more effective than trazodone: a double-blind controlled study in hospitalized patients with major depression. Int Clin Psychopharmacol. 1995 Mar;10(1):3-9.

96. Cunningham LA, Borison RL, Carman JS, Chouinard G, Crowder JE, Diamond BI, et al. A comparison of venlafaxine, trazodone, and placebo in major depression. J Clin Psychopharmacol. 1994 Apr;14(2):99-106.

97. Weisler RH, Johnston JA, Lineberry CG, Samara B, Branconnier RJ, Billow AA. Comparison of bupropion and trazodone for the treatment of major depression. J Clin Psychopharmacol. 1994 Jun;14(3):170-9.

98. FDA Center for Drug Evaluation and Research. Statistical review of NDA 21-323 (Escitalopram Oxalate). 2001 [cited; Available from: http://www.fda.gov/cder/foi/nda/2002/21-323.pdf_Lexapro_Statr.pdf

99. Ravindran AV, Guelfi JD, Lane RM, Cassano GB. Treatment of dysthymia with sertraline: a double-blind, placebo-controlled trial in dysthymic patients without major depression. J Clin Psychiatry. 2000 Nov;61(11):821-7.

100. Devanand DP, Nobler MS, Cheng J, Turret N, Pelton GH, Roose SP, et al. Randomized, double-blind, placebo-controlled trial of fluoxetine treatment for elderly patients with dysthymic disorder. Am J Geriatr Psychiatry. 2005 Jan;13(1):59-68.

101. Vanelle JM, Attar-Levy D, Poirier MF, Bouhassira M, Blin P, Olie JP. Controlled efficacy study of fluoxetine in dysthymia. Br J Psychiatry. 1997 Apr;170:345-50.

102. 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 Apr;20(2):122-8.

103. 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 May;41(5):582-92.

104. Smith D, Dempster C, Glanville J, Freemantle N, Anderson I. Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. Br J Psychiatry. 2002 May;180:396-404.

105. Gillin JC, Rapaport M, Erman MK, Winokur A, Albala BJ. A comparison of nefazodone and fluoxetine on mood and on objective, subjective, and clinician-rated measures of sleep in depressed patients: a double-blind, 8-week clinical trial. J Clin Psychiatry. 1997 May;58(5):185-92.

106. Armitage R, Yonkers K, Cole D, Rush AJ. A multicenter, double-blind comparison of the effects of nefazodone and fluoxetine on sleep architecture and quality of sleep in depressed outpatients. J Clin Psychopharmacol. 1997 Jun;17(3):161-8.

107. Doraiswamy PM, Khan ZM, Donahue RM, Richard NE. Quality of life in geriatric depression: a comparison of remitters, partial responders, and nonresponders. Am J Geriatr Psychiatry. 2001 Fall;9(4):423-8.

108. Rush AJ, Trivedi MH, Carmody TJ, Donahue RM, Houser TL, Bolden-Watson C, et al. Response in relation to baseline anxiety levels in major depressive disorder treated with bupropion sustained release or

sertraline. Neuropsychopharmacology. 2001 Jul;25(1):131-8.

109. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. JAMA. 2006 Mar 8;295(10):1152-60.

110. Kocsis JH, Zisook S, Davidson J, Shelton R, Yonkers K, Hellerstein DJ, et al. Double-blind comparison of sertraline, imipramine, and placebo in the treatment of dysthymia: psychosocial outcomes. Am J Psychiatry. 1997 Mar;154(3):390-5.

111. Thase ME, Fava M, Halbreich U, Kocsis JH, Koran L, Davidson J, et al. A placebo-controlled, randomized clinical trial comparing sertraline and imipramine for the treatment of dysthymia. Arch Gen Psychiatry. 1996 Sep;53(9):777-84.

112. Williams JW, Jr., Barrett J, Oxman T, Frank E, Katon W, Sullivan M, et al. Treatment of dysthymia and minor depression in primary care: A randomized controlled trial in older adults. JAMA. 2000 Sep 27;284(12):1519-26.

113. Barrett JE, Williams JW, Jr., Oxman TE, Frank E, Katon W, Sullivan M, et al. Treatment of dysthymia and minor depression in primary care: a randomized trial in patients aged 18 to 59 years. J Fam Pract. 2001 May;50(5):405-12.

114. Rocca P, Calvarese P, Faggiano F, Marchiaro L, Mathis F, Rivoira E, et al. Citalopram versus sertraline in late-life nonmajor clinically significant depression: a 1-year follow-up clinical trial. J Clin Psychiatry. 2005 Mar;66(3):360-9.

115. Judd LL, Rapaport MH, Yonkers KA, Rush AJ, Frank E, Thase ME, et al. Randomized, placebo-controlled trial of fluoxetine for acute treatment of minor depressive disorder. Am J Psychiatry. 2004 Oct;161(10):1864-71.

116. 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 Mar;58(3):104-7.

117. Franchini L, Gasperini M, Zanardi R, Smeraldi E. Four-year follow-up study of sertraline and fluvoxamine in long-term treatment of unipolar subjects with high recurrence rate. J Affect Disord. 2000 Jun;58(3):233-6.

118. Weihs KL, Houser TL, Batey SR, Ascher JA, Bolden-Watson C, Donahue RM, et al. Continuation phase treatment with bupropion SR effectively decreases the risk for relapse of depression. Biol Psychiatry. 2002 May 1;51(9):753-61.

119. Montgomery SA, Dunbar G. Paroxetine is better than placebo in relapse prevention and the prophylaxis of recurrent depression. Int Clin Psychopharmacol. 1993 Fall;8(3):189-95.

120. Robert P, Montgomery SA. Citalopram in doses of 20-60 mg is effective in depression relapse prevention: a placebo-controlled 6 month study. Int Clin Psychopharmacol. 1995 Mar;10 Suppl 1:29-35.

121. Rapaport MH, Bose A, Zheng H. Escitalopram continuation treatment prevents relapse of depressive episodes. J Clin Psychiatry. 2004 Jan;65(1):44-9.

122. Reimherr FW, Amsterdam JD, Quitkin FM, Rosenbaum JF, Fava M, Zajecka J, et al. Optimal length of continuation therapy in depression: a prospective assessment during long-term fluoxetine treatment. Am J Psychiatry. 1998 Sep;155(9):1247-53.

123. Schmidt ME, Fava M, Robinson JM, Judge R. The efficacy and safety of a new enteric-coated formulation of fluoxetine given once weekly during the continuation treatment of major depressive disorder. J Clin Psychiatry. 2000 Nov;61(11):851-7.

124. Thase ME, Nierenberg AA, Keller MB, Panagides J. Efficacy of mirtazapine for prevention of depressive relapse: a placebo-controlled double-blind trial of recently remitted high-risk patients. J Clin Psychiatry. 2001 Oct;62(10):782-8.

125. Feiger AD, Bielski RJ, Bremner J, Heiser JF, Trivedi M, Wilcox CS, et al. Double-blind, placebo-substitution study of nefazodone in the prevention of relapse during continuation treatment of outpatients with major depression. Int Clin Psychopharmacol. 1999 Jan;14(1):19-28.

126. Doogan DP, Caillard V. Sertraline in the prevention of depression. Br J Psychiatry. 1992 Feb;160:217-22.

127. Simon JS, Aguiar LM, Kunz NR, Lei D. Extended-release venlafaxine in relapse prevention for patients with major depressive disorder. J Psychiatr Res. 2004 May-Jun;38(3):249-57.

128. Hochstrasser B, Isaksen PM, Koponen H, Lauritzen L, Mahnert FA, Rouillon F, et al. Prophylactic effect of citalopram in unipolar, recurrent depression: placebo-controlled study of maintenance therapy. Br J Psychiatry. 2001 Apr;178:304-10.

129. Klysner R, Bent-Hansen J, Hansen HL, Lunde M, Pleidrup E, Poulsen DL, et al. Efficacy of citalopram in the prevention of recurrent depression in elderly patients: placebo-controlled study of maintenance therapy. Br J Psychiatry. 2002 Jul;181:29-35.

130. Gilaberte I, Montejo AL, de la Gandara J, Perez-Sola V, Bernardo M, Massana J, et al. Fluoxetine in the prevention of depressive recurrences: a double-blind

study. J Clin Psychopharmacol. 2001 Aug;21(4):417-24.

131. Terra JL, Montgomery SA. Fluvoxamine prevents recurrence of depression: results of a long-term, double-blind, placebo-controlled study. Int Clin Psychopharmacol. 1998 Mar;13(2):55-62.

132. Gelenberg AJ, Trivedi MH, Rush AJ, Thase ME, Howland R, Klein DN, et al. Randomized, placebo-controlled trial of nefazodone maintenance treatment in preventing recurrence in chronic depression. Biol Psychiatry. 2003 Oct 15;54(8):806-17.

133. Claghorn JL, Feighner JP. A double-blind comparison of paroxetine with imipramine in the long-term treatment of depression. J Clin Psychopharmacol. 1993 Dec;13(6 Suppl 2):23S-7S.

134. Keller MB, Kocsis JH, Thase ME, Gelenberg AJ, Rush AJ, Koran L, et al. Maintenance phase efficacy of sertraline for chronic depression: a randomized controlled trial. Jama. 1998 Nov 18;280(19):1665-72.

135. Kocsis JH, Schatzberg A, Rush AJ, Klein DN, Howland R, Gniwesch L, et al. Psychosocial outcomes following long-term, double-blind treatment of chronic depression with sertraline vs placebo. Arch Gen Psychiatry. 2002 Aug;59(8):723-8.

136. Lepine JP, Caillard V, Bisserbe JC, Troy S, Hotton JM, Boyer P. A randomized, placebo-controlled trial of sertraline for prophylactic treatment of highly recurrent major depressive disorder. Am J Psychiatry. 2004 May;161(5):836-42.

137. Wilson KC, Mottram PG, Ashworth L, Abou-Saleh MT. Older community residents with depression: long-term treatment with sertraline. Randomised, double-blind, placebo-controlled study. Br J Psychiatry. 2003 Jun;182:492-7.

138. Montgomery SA, Entsuah R, Hackett D, Kunz NR, Rudolph RL. Venlafaxine versus placebo in the preventive treatment of recurrent major depression. J Clin Psychiatry. 2004 Mar;65(3):328-36.

139. Trivedi MH, Fava M, Wisniewski SR, Thase ME, Quitkin F, Warden D, et al. Medication augmentation after the failure of SSRIs for depression. N Engl J Med. 2006 Mar 23;354(12):1243-52.

140. Baldomero EB, Ubago JG, Cercos CL, Ruiloba JV, Calvo CG, Lopez RP. Venlafaxine extended release versus conventional antidepressants in the remission of depressive disorders after previous antidepressant failure: ARGOS study. Depress Anxiety. 2005 Aug 10;22(2):68-76.

141. Poirier MF, Boyer P. Venlafaxine and paroxetine in treatment-resistant depression. Double-blind, randomised comparison. Br J Psychiatry. 1999 Jul;175:12-6.

142. Dinan TG. Efficacy and safety of weekly treatment with enteric-coated fluoxetine in patients with major depressive disorder. J Clin Psychiatry. 2001;62 Suppl 22:48-52.

143. Montgomery SA, Rasmussen JG. Citalopram 20 mg, citalopram 40 mg and placebo in the prevention of relapse of major depression. Int Clin Psychopharmacol. 1992 Jun;8:181-8.

144. Montgomery SA, Rasmussen JG, Lyby K, Connor P, Tanghoj P. Dose response relationship of citalopram 20 mg, citalopram 40 mg and placebo in the treatment of moderate and severe depression. Int Clin Psychopharmacol. 1992 Jun;6 Suppl 5:65-70.

145. Michelson D, Amsterdam JD, Quitkin FM, Reimherr FW, Rosenbaum JF, Zajecka J, et al. Changes in weight during a 1-year trial of fluoxetine. Am J Psychiatry. 1999 Aug;156(8):1170-6.

146. Rush AJ, Trivedi MH, Wisniewski SR, Stewart JW, Nierenberg AA, Thase ME, et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. N Engl J Med. 2006 Mar 23;354(12):1231-42.

147. Gaynes BN, Rush AJ, Trivedi M, Wisniewski SR, Balasubramani GK, Spencer DC, et al. A direct comparison of presenting characteristics of depressed outpatients from primary vs. specialty care settings: preliminary findings from the STAR*D clinical trial. Gen Hosp Psychiatry. 2005 Mar-Apr;27(2):87-96.

148. Trivedi MH, Rush AJ, Carmody TJ, Donahue RM, Bolden-Watson C, Houser TL, et al. Do bupropion SR and sertraline differ in their effects on anxiety in depressed patients? J Clin Psychiatry. 2001 Oct;62(10):776-81.

149. Flament MF, Lane RM, Zhu R, Ying Z. Predictors of an acute antidepressant response to fluoxetine and sertraline. Int Clin Psychopharmacol. 1999 Sep;14(5):259-75.

150. Fava M, Rosenbaum JF, Hoog SL, Tepner RG, Kopp JB, Nilsson ME. Fluoxetine versus sertraline and paroxetine in major depression: tolerability and efficacy in anxious depression. J Affect Disord. 2000 Aug;59(2):119-26.

151. Lader M, Andersen HF, Baekdal T. The effect of escitalopram on sleep problems in depressed patients. Hum Psychopharmacol. 2005 Jul;20(5):349-54.

152. Clerc GE, Ruimy P, Verdeau-Palles J. A double-blind comparison of venlafaxine and fluoxetine in patients hospitalized for major depression and melancholia.

The Venlafaxine French Inpatient Study Group. Int Clin Psychopharmacol. 1994 Sep;9(3):139-43.

153. Goldstein DJ, Lu Y, Detke MJ, Wiltse C, Mallinckrodt C, Demitrack MA. Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. J Clin Psychopharmacol. 2004 Aug;24(4):389-99.

154. Brannan SK, Mallinckrodt CH, Brown EB, Wohlreich MM, Watkin JG, Schatzberg AF. Duloxetine 60 mg once-daily in the treatment of painful physical symptoms in patients with major depressive disorder. J Psychiatr Res. 2005 Jan;39(1):43-53.

155. Detke MJ, Lu Y, Goldstein DJ, Hayes JR, Demitrack MA. Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. J Clin Psychiatry. 2002 Apr;63(4):308-15.

156. Detke MJ, Lu Y, Goldstein DJ, McNamara RK, Demitrack MA. Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. J Psychiatr Res. 2002 Nov-Dec;36(6):383-90.

157. Joliat MJ, Schmidt ME, Fava M, Zhang S, Michelson D, Trapp NJ, et al. Long-term treatment outcomes of depression with associated anxiety: efficacy of continuation treatment with fluoxetine. J Clin Psychiatry. 2004 Mar;65(3):373-8.

158. Khan A, Upton GV, Rudolph RL, Entsuah R, Leventer SM. The use of venlafaxine in the treatment of major depression and major depression associated with anxiety: a dose-response study. Venlafaxine Investigator Study Group. J Clin Psychopharmacol. 1998 Feb;18(1):19-25.

159. Mallinckrodt CH, Watkin JG, Liu C, Wohlreich MM, Raskin J. Duloxetine in the treatment of Major Depressive Disorder: a comparison of efficacy in patients with and without melancholic features. BMC Psychiatry. 2005 Jan 4;5(1):1.

160. Mackay FJ, Dunn NR, Wilton LV, Pearce GL, Freemantle SN, Mann RD. A comparison of fluvoxamine, fluoxetine, sertraline and paroxetine examined by observational cohort studies. Pharmacoepidemiol Drug Saf. 1997 Jul;6(4):235-46.

161. Mackay FR, Dunn NR, Martin RM, Pearce GL, Freemantle SN, Mann RD. Newer antidepressants: a comparison of tolerability in general practice. Br J Gen Pract. 1999 Nov;49(448):892-6.

162. Meijer WE, Heerdink ER, van Eijk JT, Leufkens HG. Adverse events in users of sertraline: results from an observational study in psychiatric practice in The Netherlands. Pharmacoepidemiol Drug Saf. 2002 Dec;11(8):655-62.

163. Brambilla P, Cipriani A, Hotopf M, Barbui C. Side-effect profile of fluoxetine in comparison with other SSRIs, tricyclic and newer antidepressants: a meta-analysis of clinical trial data. Pharmacopsychiatry. 2005 Mar;38(2):69-77.

164. Greist J, McNamara RK, Mallinckrodt CH, Rayamajhi JN, Raskin J. Incidence and duration of antidepressant-induced nausea: duloxetine compared with paroxetine and fluoxetine. Clin Ther. 2004 Sep;26(9):1446-55.

165. Mackay FJ, Dunn NR, Mann RD. Antidepressants and the serotonin syndrome in general practice. Br J Gen Pract. 1999 Nov;49(448):871-4.

166. 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 Apr;24(4):662-72.

167. Goldstein DJ, Hamilton SH, Masica DN, Beasley CM, Jr. Fluoxetine in medically stable, depressed geriatric patients: effects on weight. J Clin Psychopharmacol. 1997 Oct;17(5):365-9.

168. Harto NE, Spera KF, Branconnier RJ. Fluoxetine-induced reduction of body mass in patients with major depressive disorder. Psychopharmacol Bull. 1988;24(2):220-3.

169. Anonymous. Report of the CSM expert working group on the safety of selective serotonin reuptake inhibitor antidepressants. 2004 [cited; Available from: http://www.mhra.gov.uk/home/groups/pl-p/documents/drugsafetymessage/con019472.pdf

170. Judge R, Parry MG, Quail D, Jacobson JG. Discontinuation symptoms: comparison of brief interruption in fluoxetine and paroxetine treatment. Int Clin Psychopharmacol. 2002 Sep;17(5):217-25.

171. Perahia DG, Kajdasz DK, Desaiah D, Haddad PM. Symptoms following abrupt discontinuation of duloxetine treatment in patients with major depressive disorder. J Affect Disord. 2005 Dec;89(1-3):207-12.

172. Zajecka J, Fawcett J, Amsterdam J, Quitkin F, Reimherr F, Rosenbaum J, et al. Safety of abrupt discontinuation of fluoxetine: a randomized, placebo-controlled study. J Clin Psychopharmacol. 1998 Jun;18(3):193-7.

173. Pedersen AG. Escitalopram and suicidality in adult depression and anxiety. Int Clin Psychopharmacol. 2005 May;20(3):139-43.

174. 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. Br Med J. 2005 Feb 19;330(7488):396.

175. 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 Apr;160(4):790-2.

176. 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. Br Med J. 2005 Feb 19;330(7488):385.

177. 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. Br Med J. 2005 Feb 19;330(7488):389.

178. Didham RC, McConnell DW, Blair HJ, Reith DM. Suicide and self-harm following prescription of SSRIs and other antidepressants: confounding by indication. Br J Clin Pharmacol. 2005 Nov;60(5):519-25.

179. Jick H, Ulcickas M, Dean A. Comparison of frequencies of suicidal tendencies among patients receiving fluoxetine, lofepramine, mianserin, or trazodone. Pharmacotherapy. 1992;12(6):451-4.

180. Jick H, Kaye JA, Jick SS. Antidepressants and the risk of suicidal behaviors. JAMA. 2004 Jul 21;292(3):338-43.

181. Jick SS, Dean AD, Jick H. Antidepressants and suicide. Br Med J. 1995 Jan 28;310(6974):215-8.

182. Aursnes I, Tvete IF, Gaasemyr J, Natvig B. Suicide attempts in clinical trials with paroxetine randomised against placebo. BMC Med. 2005 Aug 22;3:14.

183. Lopez-Iibor. Reduced suicidality with paroxetine. European Psychiatry. 1993;8(Suppl 1):17S-9S.

184. 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 Apr;63(4):357-66.

185. Delgado PL, Brannan SK, Mallinckrodt CH, Tran PV, McNamara RK, Wang F, et al. Sexual functioning assessed in 4 double-blind placebo- and paroxetine-controlled trials of duloxetine for major depressive disorder. J Clin Psychiatry. 2005 Jun;66(6):686-92.

186. Ekselius L, von Knorring L. Effect on sexual function of long-term treatment with selective serotonin reuptake inhibitors in depressed patients treated in primary care. J Clin Psychopharmacol. 2001 Apr;21(2):154-60.

187. Ferguson JM, Shrivastava RK, Stahl SM, Hartford JT, Borian F, Ieni J, et al. Reemergence of sexual dysfunction in patients with major depressive disorder: double-blind comparison of nefazodone and sertraline. J Clin Psychiatry. 2001 Jan;62(1):24-9.

188. Kennedy SH, Eisfeld BS, Dickens SE, Bacchiochi JR, Bagby RM. Antidepressant-induced sexual dysfunction during treatment with moclobemide, paroxetine, sertraline, and venlafaxine. J Clin Psychiatry. 2000 Apr;61(4):276-81.

189. Landen M, Hogberg P, Thase ME. Incidence of sexual side effects in refractory depression during treatment with citalopram or paroxetine. J Clin Psychiatry. 2005 Jan;66(1):100-6.

190. Montejo AL, Llorca G, Izquierdo JA, Rico-Villademoros F. Incidence of sexual dysfunction associated with antidepressant agents: a prospective multicenter study of 1022 outpatients. Spanish Working Group for the Study of Psychotropic-Related Sexual Dysfunction. J Clin Psychiatry. 2001;62 Suppl 3:10-21.

191. Nieuwstraten CE, Dolovich LR. Bupropion versus selective serotonin-reuptake inhibitors for treatment of depression. Ann Pharmacother. 2001 Dec;35(12):1608-13.

192. Philipp M, Tiller JW, Baier D, Kohnen R. Comparison of moclobemide with selective serotonin reuptake inhibitors (SSRIs) on sexual function in depressed adults. The Australian and German Study Groups. Eur Neuropsychopharmacol. 2000 Sep;10(5):305-14.

193. 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 Jul;59(7):366-73.

194. 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 Nov;52(11):450-6.

195. Whyte IM, Dawson AH, Buckley NA. Relative toxicity of venlafaxine and selective serotonin reuptake inhibitors in overdose compared to tricyclic antidepressants. Qjm. 2003 May;96(5):369-74.

196. Thase ME. Effects of venlafaxine on blood pressure: a meta-analysis of original data from 3744 depressed patients. J Clin Psychiatry. 1998 Oct;59(10):502-8.

197. Thase ME, Tran PV, Wiltse C, Pangallo BA, Mallinckrodt C, Detke MJ. Cardiovascular profile of duloxetine, a dual reuptake inhibitor of serotonin and norepinephrine. J Clin Psychopharmacol. 2005 Apr;25(2):132-40.

198. Buckley NA, McManus PR. Fatal toxicity of serotoninergic and other antidepressant drugs: analysis of United Kingdom mortality data. Br Med J. 2002 Dec 7;325(7376):1332-3.

199. Coogan PF, Palmer JR, Strom BL, Rosenberg L. Use of selective serotonin reuptake inhibitors and the risk of breast cancer. Am J Epidemiol. 2005 Nov 1;162(9):835-8.

200. Kirby D, Harrigan S, Ames D. Hyponatraemia in elderly psychiatric patients treated with Selective Serotonin Reuptake Inhibitors and venlafaxine: a retrospective controlled study in an inpatient unit. Int J Geriatr Psychiatry. 2002 Mar;17(3):231-7.

201. Thapa PB, Gideon P, Cost TW, Milam AB, Ray WA. Antidepressants and the risk of falls among nursing home residents. N Engl J Med. 1998 Sep 24;339(13):875-82.

202. Paroxetine Adult Suicidality Analysis: Major Depressive Disorder and Non-Major Depressive Disorder. 2006 [cited; Available from: http://www.gsk.com/media/paroxetine/briefing_doc.pdf.

203. Liu BA, Mittmann N, Knowles SR, Shear NH. Hyponatremia and the syndrome of innappropriate secretion of antidiuretic hormone associated with the use of selective serotonin reuptake inhibitors: a review of spontaneous reports. CMAJ. 1996;155(5):519-27.

204. Stewart DE. Hepatic adverse reactions associated with nefazodone. Can J Psychiatry. 2002;47(4):375-7.

205. Mason PJ, Morris VA, Balcezak TJ. Serotonin syndrome. Presentation of 2 cases and review of the literature. Medicine (Baltimore). 2000 Jul;79(4):201-9.

206. Keene MS, Eaddy MT, Mauch RP, Regan TS, Shah M, Chiao E. Differences in compliance patterns across the selective serotonin reuptake inhibitors (SSRIs). Curr Med Res Opin. 2005 Oct;21(10):1651-8.

207. Miller LG, Hays RD. Adherence to combination antiretroviral therapy: synthesis of the literature and clinical implications. AIDS Read. 2000 Mar;10(3):177-85.

208. Hernandez-Avila CA, Modesto-Lowe V, Feinn R, Kranzler HR. Nefazodone treatment of comorbid alcohol dependence and major depression. Alcohol Clin Exp Res. 2004 Mar;28(3):433-40.

209. Krishnan KR, Doraiswamy PM, Clary CM. Clinical and treatment response characteristics of late-life depression associated with vascular disease: a pooled analysis of two multicenter trials with sertraline. Prog Neuropsychopharmacol Biol Psychiatry. 2001 Feb;25(2):347-61.

210. Moak DH, Anton RF, Latham PK, Voronin KE, Waid RL, Durazo-Arvizu R. Sertraline and cognitive behavioral therapy for depressed alcoholics: results of a placebo-controlled trial. J Clin Psychopharmacol. 2003 Dec;23(6):553-62.

211. Nyth AL, Gottfries CG, Lyby K, Smedegaard-Andersen L, Gylding-Sabroe J, Kristensen M, et al. A controlled multicenter clinical study of citalopram and placebo in elderly depressed patients with and without concomitant dementia. Acta Psychiatr Scand. 1992 Aug;86(2):138-45.

212. Petrakis I, Carroll KM, Nich C, Gordon L, Kosten T, Rounsaville B. Fluoxetine treatment of depressive disorders in methadone-maintained opioid addicts. Drug Alcohol Depend. 1998 May 1;50(3):221-6.

213. Rapaport MH, Schneider LS, Dunner DL, Davies JT, Pitts CD. Efficacy of controlled-release paroxetine in the treatment of late-life depression. J Clin Psychiatry. 2003 Sep;64(9):1065-74.

214. Geretsegger C, Bohmer F, Ludwig M. Paroxetine in the elderly depressed patient: randomized comparison with fluoxetine of efficacy, cognitive and behavioural effects. Int Clin Psychopharmacol. 1994 Spring;9(1):25-9.

215. 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 Aug;64(8):875-82.

216. Thase ME, Entsuah R, Cantillon M, Kornstein SG. Relative Antidepressant Efficacy of Venlafaxine and SSRIs: Sex-Age Interactions. J Womens Health (Larchmt). 2005 Sep;14(7):609-16.

217. 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 Nov;62(11):869-77.

218. Burt VK, Wohlreich MM, Mallinckrodt CH, Detke MJ, Watkin JG, Stewart DE. Duloxetine for the treatment of major depressive disorder in women ages 40 to 55 years. Psychosomatics. 2005 Jul-Aug;46(4):345-54.

219. Roose SP, Sackeim HA, Krishnan KR, Pollock BG, Alexopoulos G, Lavretsky H, et al. Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. Am J Psychiatry. 2004 Nov;161(11):2050-9.

220. Schneider LS, Nelson JC, Clary CM, Newhouse P, Krishnan KR, Shiovitz T, et al. An 8-week multicenter, parallel-group, double-blind, placebo-

controlled study of sertraline in elderly outpatients with major depression. Am J Psychiatry. 2003 Jul;160(7):1277-85.

221. Sheikh JI, Cassidy EL, Doraiswamy PM, Salomon RM, Hornig M, Holland PJ, et al. Efficacy, safety, and tolerability of sertraline in patients with late-life depression and comorbid medical illness. J Am Geriatr Soc. 2004 Jan;52(1):86-92.

222. Tollefson GD, Holman SL. Analysis of the Hamilton Depression Rating Scale factors from a double-blind, placebo-controlled trial of fluoxetine in geriatric major depression. Int Clin Psychopharmacol. 1993 Winter;8(4):253-9.

223. Tollefson GD, Bosomworth JC, Heiligenstein JH, Potvin JH, Holman S. A double-blind, placebo-controlled clinical trial of fluoxetine in geriatric patients with major depression. The Fluoxetine Collaborative Study Group. Int Psychogeriatr. 1995 Spring;7(1):89-104.

224. Small GW, Birkett M, Meyers BS, Koran LM, Bystritsky A, Nemeroff CB. Impact of physical illness on quality of life and antidepressant response in geriatric major depression. Fluoxetine Collaborative Study Group. J Am Geriatr Soc. 1996 Oct;44(10):1220-5.

225. Wagner GJ, Maguen S, Rabkin JG. Ethnic differences in response to fluoxetine in a controlled trial with depressed HIV-positive patients. Psychiatr Serv. 1998 Feb;49(2):239-40.

226. Ferrando SJ, Goldman JD, Charness WE. Selective serotonin reuptake inhibitor treatment of depression in symptomatic HIV infection and AIDS. Improvements in affective and somatic symptoms. Gen Hosp Psychiatry. 1997 Mar;19(2):89-97.

227. Rabkin JG, Wagner GJ, McElhiney MC, Rabkin R, Lin SH. Testosterone versus fluoxetine for depression and fatigue in HIV/AIDS: a placebo-controlled trial. J Clin Psychopharmacol. 2004 Aug;24(4):379-85.

228. Rabkin JG, Wagner GJ, Rabkin R. Fluoxetine treatment for depression in patients with HIV and AIDS: a randomized, placebo-controlled trial. Am J Psychiatry. 1999 Jan;156(1):101-7.

229. Gual A, Balcells M, Torres M, Madrigal M, Diez T, Serrano L. Sertraline for the prevention of relapse in detoxicated alcohol dependent patients with a comorbid depressive disorder: a randomized controlled trial. Alcohol Alcohol. 2003 Nov-Dec;38(6):619-25.

230. Lyketsos CG, DelCampo L, Steinberg M, Miles Q, Steele CD, Munro C, et al. Treating depression in Alzheimer disease: efficacy and safety of sertraline therapy, and the benefits of depression reduction: the DIADS. Arch Gen Psychiatry. 2003 Jul;60(7):737-46.

231. Magai C, Kennedy G, Cohen CI, Gomberg D. A controlled clinical trial of sertraline in the treatment of depression in nursing home patients with late-stage Alzheimer's disease. Am J Geriatr Psychiatry. 2000 Winter;8(1):66-74.

232. Roscoe JA, Morrow GR, Hickok JT, Mustian KM, Griggs JJ, Matteson SE, et al. Effect of paroxetine hydrochloride (Paxil) on fatigue and depression in breast cancer patients receiving chemotherapy. Breast Cancer Res Treat. 2005 Feb;89(3):243-9.

233. Bush DE, Ziegelstein RC, Patel UV, Thombs BD, Ford DE, Fauerbach JA, et al. Post-myocardial infarction depression. Evid Rep Technol Assess (Summ). 2005 May(123):1-8.

234. Glassman AH, O'Connor CM, Califf RM, Swedberg K, Schwartz P, Bigger JT, Jr., et al. Sertraline treatment of major depression in patients with acute MI or unstable angina. JAMA. 2002 Aug 14;288(6):701-9.

235. Strik JJ, Honig A, Lousberg R, Lousberg AH, Cheriex EC, Tuynman-Qua HG, et al. Efficacy and safety of fluoxetine in the treatment of patients with major depression after first myocardial infarction: findings from a double-blind, placebo-controlled trial. Psychosom Med. 2000 Nov-Dec;62(6):783-9.

236. Blumenfield M, Levy NB, Spinowitz B, Charytan C, Beasley CM, Jr., Dubey AK, et al. Fluoxetine in depressed patients on dialysis. Int J Psychiatry Med. 1997;27(1):71-80.

237. Andersen G, Vestergaard K, Lauritzen L. Effective treatment of poststroke depression with the selective serotonin reuptake inhibitor citalopram. Stroke. 1994 Jun;25(6):1099-104.

238. Murray V, von Arbin M, Bartfai A, Berggren AL, Landtblom AM, Lundmark J, et al. Double-blind comparison of sertraline and placebo in stroke patients with minor depression and less severe major depression. J Clin Psychiatry. 2005 Jun;66(6):708-16.

239. Schmitz JM, Averill P, Stotts AL, Moeller FG, Rhoades HM, Grabowski J. Fluoxetine treatment of cocaine-dependent patients with major depressive disorder. Drug Alcohol Depend. 2001 Aug 1;63(3):207-14.

240. Bondareff W, Alpert M, Friedhoff AJ, Richter EM, Clary CM, Batzar E. Comparison of sertraline and nortriptyline in the treatment of major depressive disorder in late life. Am J Psychiatry. 2000 May;157(5):729-36.

241. Hansen RA, Gartlehner G, Lohr KN, Gaynes BN, Carey TS. Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. Ann Intern Med. 2005 Sep 20;143(6):415-26.

242. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth edition, Text Revision. Washington, DC. 2000.

243. Fava M, Alpert JE, Carmin CN, Wisniewski SR, Trivedi MH, Biggs MM, et al. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR*D. Psychol Med. 2004 Oct;34(7):1299-308.

244. Gaynes BN, Magruder KM, Burns BJ, Wagner HR, Yarnall KS, Broadhead WE. Does a coexisting anxiety disorder predict persistence of depressive illness in primary care patients with major depression? Gen Hosp Psychiatry. 1999 May-Jun;21(3):158-67.

245. Russell JM, Koran LM, Rush J, Hirschfeld RM, Harrison W, Friedman ES, et al. Effect of concurrent anxiety on response to sertraline and imipramine in patients with chronic depression. Depress Anxiety. 2001;13(1):18-27.

246. Fava M, Uebelacker LA, Alpert JE, Nierenberg AA, Pava JA, Rosenbaum JF. Major depressive subtypes and treatment response. Biol Psychiatry. 1997 Oct 1;42(7):568-76.

247. Panzer MJ. Are SSRIs really more effective for anxious depression? Ann Clin Psychiatry. 2005 Jan-Mar;17(1):23-9.

248. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. Arch Intern Med. 2003 Nov 10;163(20):2433-45.

249. Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? Acad Emerg Med. 1998 Nov;5(11):1086-90.

250. Kelly AM. The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. Emerg Med J. 2001 May;18(3):205-7.

251. ten Klooster PM, Drossaers-Bakker KW, Taal E, van de Laar MA. Patient-perceived satisfactory improvement (PPSI): interpreting meaningful change in pain from the patient's perspective. Pain. 2006 Mar;121(1-2):151-7.

252. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. Am J Psychiatry. 2006 Nov;163(11):1905-17.