Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder

Hansen R A, Gartlehner G, Lohr K N, Gaynes B N, Carey T S

CRD summary
This was a well-conducted review of second-generation antidepressants in the treatment of major depressive disorder. No substantial differences were found in the efficacy of the drugs examined. However, statistical synthesis was not possible for many drug comparisons because the quantity and quality of the evidence was limited. Good-quality trials with a focus upon clinically relevant outcomes were recommended.

Authors’ objectives
To compare data on the efficacy, effectiveness and tolerability of second-generation antidepressants in the treatment of major depressive disorder.

Searching
MEDLINE, EMBASE, PsycLIT, the Cochrane Library and International Pharmaceutical Abstracts were searched from 1980 to February 2005. Reference lists of relevant review articles and letters to the editor were scrutinised, and pharmaceutical manufacturers were consulted. Attempts were made to retrieve unpublished data from the U.S. Food and Drug Administration, but these were unsuccessful.

Study selection
Study designs of evaluations included in the review
For the assessment of efficacy or effectiveness, RCTs of at least 6 weeks' duration were eligible for inclusion. For the assessment of safety or tolerability, RCTs of at least 6 weeks' duration, placebo-controlled trials designed specifically to assess adverse events, and observational studies with at least 100 participants and covering at least 1 year were eligible for inclusion. RCTs of effectiveness had to have a minimum of 3 months' follow-up, use minimal inclusion or exclusion criteria, and have an adequate sample size to determine a minimally important difference on a health-related quality-of-life instrument.

Specific interventions included in the review
Studies of commonly prescribed second-generation antidepressants were eligible for inclusion. The included drugs comprised six selective serotonin re-uptake inhibitors (SSRIs; citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine and sertraline) and four other second-generation antidepressants (bupropion, duloxetine, mirtazapine and venlafaxine).

Participants included in the review
Studies of paediatric or adult out-patients with major depressive disorder were eligible for inclusion. For the assessment of effectiveness, randomised controlled trials (RCTs) were those conducted in a primary care setting. The majority of the included participants were younger than 60 years; in 13% of trials the participants were 60 years or older, and in 7% of trials the participants were children or adolescents (younger than 18 years).

Outcomes assessed in the review
For the assessment of efficacy or effectiveness, RCTs had to assess health outcomes rather than intermediate outcomes. The primary outcome measure was treatment response, which was defined as a 50% or greater improvement in the Hamilton Rating Scale for Depression (HAM-D) or the Montgomery-Asberg Depression Rating Scale. Speed of response, quality of life, remission, relapse, functional capacity, hospitalisation and adverse events were also measured.

How were decisions on the relevance of primary studies made?
Two reviewers independently reviewed the titles and abstracts. It was unclear how any disagreements were resolved. [A: Any disagreements were resolved by discussion and consultation with a third reviewer.]
Assessment of study quality
Validity was assessed using criteria from the U.S. Preventive Services Task Force (rated good, fair, or poor) and the Centre for Reviews and Dissemination. The criteria included randomisation, allocation concealment, similarity of comparison groups, use of intention-to-treat analysis, and loss to follow-up. One reviewer conducted the quality assessment and a second reviewer verified the completeness of the assessment. It was unclear how any discrepancies were resolved. [A: Any disagreements were resolved by discussion and consultation with a third reviewer.]

Data extraction
One reviewer extracted the data and a second reviewer verified the completeness of the extraction. It was unclear how any discrepancies were resolved. [A: Any disagreements were resolved by discussion and consultation with a third reviewer.] Data on the percentage improvements in response were extracted, on an intention-to-treat basis where possible. The mean incidence and 95% confidence intervals (CIs) were calculated for adverse events.

Methods of synthesis
How were the studies combined?
The studies were summarised qualitatively and, where more than 3 RCTS had compared the same treatment, a meta-analysis was used to combine these studies. In homogeneous studies, random-effects and fixed-effect models was used. The results of the former were reported on the basis that they represented a more conservative estimate. The presence of publication bias was estimated using funnel plots, the Begg adjusted rank correlation test and the Egger regression approach.

How were differences between studies investigated?
Heterogeneity of treatment effects in each meta-analysis was assessed using the I-squared statistic. When treatment effects differed, the authors explored potential reasons for such differences.

Results of the review
Seventy studies were included in the review, of which 46 RCTs (n=11,553) assessed efficacy or effectiveness. Twenty-two of these (n=5,333) compared SSRIs with other SSRIs and 24 (n=6,220) compared an SSRI with another second-generation antidepressant. Twenty-four observational or placebo-controlled studies were used in the analysis of adverse events (number of participants not given).

SSRI compared with another SSRI.
Twenty-one studies were considered to be of a fair quality and one study was rated as good. Three studies were classed as effectiveness trials. The majority of the efficacy trials and all effectiveness trials reported no statistically significant differences on any outcome measure. Two meta-analyses were conducted. Six studies were pooled (774 patients) to compare fluoxetine with paroxetine, demonstrating no statistically significant differences between the drugs (relative benefit 1.09, 95% CI: 0.97, 1.21). No major publication bias was found in the Begg adjusted rank order correlation test (Kendall tau 0.3; P=0.47) and the Egger regression approach (intercept 2.107, 95% CI: -0.237, 4.450). Five studies were pooled (1,190 patients) to compare fluoxetine with sertraline, demonstrating a modest statistically significant treatment benefit for sertraline (relative benefit 1.10, 95% CI: 1.01, 1.20). Again, no publication bias was observed in the Begg adjusted rank order correlation test (Kendall tau 0.2; P=0.82) or the Egger regression approach (intercept 0.799, 95% CI: -6.069, 7.668).

In the 9 studies that measured health-related quality of life, good to fair evidence suggested that SSRIs did not differ substantially in terms of improvements on this outcome. Sleep quality was reported to be better in fluvoxamine-treated patients in the one trial that compared this drug with fluoxetine.

The majority of trials found no difference in terms of speed of the patient’s response to SSRIs.

SSRIs compared with other second-generation antidepressants.
All 24 trials were classified as efficacy studies. Twenty-one trials were recorded as being of a fair quality and 3 were rated good. The majority of the trials demonstrated similar efficacy between the drugs. There was some evidence in favour of a higher response rate for venlafaxine-treated patients; however, few individual trials reached statistical significance. One meta-analysis was conducted. Six studies were pooled (1,340 patients) to compare venlafaxine with fluoxetine, demonstrating a modest statistically significant treatment benefit for venlafaxine (relative benefit 1.12, 95% CI: 1.02, 1.23) on the HAM-D scale. There was no observed publication bias in the Begg adjusted rank order correlation test (Kendall tau 0.067; P>0.99) or Egger regression approach (intercept 0.141 (95% CI: -4.158, 4.440).

In the 5 trials that assessed quality of life or health-related functional capacity, no significant differences in overall quality of life were reported. Mirtazapine-treated patients experienced better sleep quality in one further trial that compared this drug with sertraline.

No statistically significant differences in response rates were noted in the 20 trials that measured this outcome.

The overall incidence of adverse events was similar amongst the drugs examined. However, there were noticeable differences in the nature of these adverse events. The authors advised cautious interpretation of these results given the differences in the assessment and reporting techniques.

Authors' conclusions
Second-generation antidepressants do not differ substantially in the treatment of major depressive disorder.

CRD commentary
The review question was clear and was supported by a thorough search strategy to identify relevant studies. The authors acknowledged that industry sponsorship was a potential major threat to the reliability of the findings, and publication bias was appropriately explored. Attempts were made to minimise possible biases in the review process. Validity was assessed using appropriate criteria, and the results of this assessment were used to highlight the existence of a very limited amount of good-quality evidence. Heterogeneity was assessed and possible reasons for conflicting results were explored. However, the inability to use many of the comparisons in a meta-analysis was regarded as a significant study limitation. This was a well-conducted systematic review and the authors' conclusion is a reliable reflection of the evidence presented.

Implications of the review for practice and research
Practice: The authors stated that the prescription of second-generation antidepressants should be based on the individual patient, patient preference, expected side-effects, and cost.

Research: The authors recommended large, high-quality effectiveness trials with head-to-head comparisons that focus on clinically relevant outcomes. In addition, adverse events should be defined a priori and assessed using consistent measurement and reporting techniques.

Funding
Center for Evidence-Based Policy, Oregon Health and Science University; 12 state Medicaid programmes; Canadian Coordinating Office for Health Technology Assessment; California HealthCare Foundation.

Bibliographic details

PubMedID
16172440
Original Paper URL
http://www.annals.org/cgi/content/full/143/6/415

Indexing Status
Subject indexing assigned by NLM

MeSH
Antidepressive Agents, Second-Generation /adverse effects /therapeutic use; Depressive Disorder, Major /drug therapy; Humans; Quality of Life; Randomized Controlled Trials as Topic; Serotonin Uptake Inhibitors /adverse effects /therapeutic use

AccessionNumber
12005008495

Date bibliographic record published
31/03/2006

Date abstract record published
31/03/2006

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.