How well do antidepressants work in older people: a systematic review of number needed to treat

Katona C, Livingston G

Authors' objectives
To directly compare the efficacy and safety of antidepressants in older people using the number-needed-to-treat (NNT) and the number-needed-to-harm (NNH).

Searching
MEDLINE (from 1966 to 1999) and EMBASE (from 1994 to 1999) were searched. The authors stated that full details of the search were reported in a previous review (see Other Publications of Related Interest). Manufacturers of all antidepressants were contacted for data held on antidepressant trials in older people.

Study selection
Double-blind randomised controlled trials (RCTs) that lasted at least 4 weeks were eligible for inclusion if they reported the number of patients in each treatment group. Studies were only included if they presented sufficient data for calculation of the NNT. Case reports and open studies were excluded.

Specific interventions included in the review
Studies of antidepressants were eligible. The included studies compared the following antidepressants with placebo or each other: sertraline, amitriptyline, paroxetine, fluoxetine, bupropion, nortriptyline, moclobemide, fluvoxamine, venlafaxine, mirtazapine, citalopram, mianserin, clomipramine, dothiepin, trazodone, imipramine, and nomifensine.

Participants included in the review
Studies of older people (age 60 and over) who had been diagnosed with a major depressive disorder or unipolar depression, using defined criteria, were eligible for inclusion.

Outcomes assessed in the review
Studies that presented measures of efficacy and the percentage of responders were eligible for inclusion. The included studies assessed efficacy using the Hamilton Depression Rating Scale (HAMD), the Montgomery Asberg Depression Rating Scale (MADRS), or the Clinical Global Impression (CGI) measure. The HAMD classified outcomes as very much improved, improved, less than 10, response, remission, and 50% reduction. Using the MADRS, the outcomes were classified as less than 50%, marked reduction, less than 12, 50% reduction, and improvement. The CGI classified outcomes as improved, very much improved, very much or much improvement, and normal to mildly ill. The review also assessed adverse effects, which were measured in the included studies on the basis of premature cessation of treatment, side-effects, and discontinuation due to side-effects.

How were decisions on the relevance of primary studies made?
The authors did not state how the papers were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
The authors did not state that they assessed validity.

Data extraction
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

The data extracted for each study included the treatments compared, the number of patients per treatment arm, the
percentage of responders per treatment arm for each measure (CGI, MADRS and HAMD), and the percentage of
patients with adverse effects per treatment arm. The NNT and NNH, together with the respective 95% confidence
intervals (CIs), were calculated for each outcome in each study.

Methods of synthesis
How were the studies combined?
The results from the individual studies were tabulated but were not synthesised. The NNT and 95% CI were tabulated
separately for placebo-controlled studies and for studies comparing antidepressants with each other. The NNH and 95%
CI were also tabulated. The NNT and NNH were compared informally across two trials.

How were differences between studies investigated?
Differences between the studies were not investigated

Results of the review
Eight RCTs (1,154 patients) compared antidepressants with placebo. Fourteen RCTs (1,778 patients) compared
antidepressants with each other.

The authors identified two RCTs published after the previous review that had compared different antidepressants (see
Other Publications of Related Interest).

Antidepressants versus placebo: most placebo-controlled studies showed that antidepressants were more effective than
placebo. The NNTs were generally comparatively small. Apart from fluoxetine and moclobemide, all the
antidepressants had an NNT of 5 or less compared with placebo. The NNT for fluoxetine (1 RCT, 671 patients) was 32
for response assessed using HAMD, and 8 for HAMD response and remission. The NNT for moclobemide (1 RCT, 71
patients) was 7 for HAMD less than 10.

Antidepressants versus each other: only one study found a significant difference in NNT between two antidepressants.
This RCT (106 patients) showed that paroxetine significantly increased the likelihood of a greater than 50% reduction
in HAMD at 6 weeks compared with fluoxetine (37% versus 16%); the NNT was 5 (95% CI: 3, 22).

In most studies the CIs were very wide, suggesting that the sample size was too small to detect a significant difference
between treatments.

One study (365 patients) suggested that citalopram and amitriptyline were equal in efficacy (percentage with marked
reduction in MADRS was 53.6% versus 53.2%; NNT 500). The studies showed a trend in favour of serotonin re-uptake
inhibitors (SSRIs) and venlafaxine compared with tricyclics.

Adverse effects (14 RCTs including 4 placebo-controlled RCTs of different drugs): the CIs were wide and most of the
upper limits of 95% CIs included infinity. The studies showed that all antidepressants increased adverse effects in
comparison with placebo. The studies suggested a moderate reduction in adverse effects for SSRIs compared with
tricyclics. Paroxetine significantly reduced adverse effects compared with amitriptyline (NNH 3, 95% CI: 2, 15).
Venlafaxine reduced adverse effects compared with clomipramine or trazodone (1 RCT with 3 treatment arms, 170
patients); the NNH was 5 (95% CI: 3, 28) for clomipramine and 4 (95% CI: 2, 13) for trazodone.

Authors’ conclusions
SSRIs (except possibly fluoxetine) and venlafaxine are more effective than other antidepressants in the treatment of
depression in older people. The authors also concluded that presenting results as the NNT and NNH can aid decision-
making in clinical practice, but that many studies did not report adequate data to allow an estimation of the NNT and
that most studies directly comparing antidepressants were underpowered.

CRD commentary
The review question was clear in terms of the study design, participants, intervention and outcomes, although the
inclusion criteria were broadly defined in terms of outcomes. Several relevant sources were searched, the search terms were stated, and attempts were made to obtain unpublished data from drug manufacturers. It was not stated whether any language limitations had been applied. The methods used to select the studies, assess validity and extract the data were not described; hence, any efforts made to reduce errors and bias cannot be judged. Only double-blind RCTs were included; other aspects of validity were not assessed.

Only minimal details of the included studies were tabulated. The information on drug dosage, duration of treatment, number of outcomes assessed in each study, number of drop-outs and cointerventions was not presented consistently. It was not stated whether the data were extracted on an intention-to-treat basis and the effect of drop-outs on the results was not explored. There was no synthesis of the evidence supporting the superiority of SSRIs and venlafaxine over tricyclics. The authors discussed some of the limitations of the review. For example, the exclusion of many studies because of the lack of suitable data for estimating the NNT, and the small size of RCTs comparing different antidepressants. Placebo-controlled studies were also generally small and specific antidepressants were studied in three RCTs at most. The evidence presented suggests that there were insufficient data to directly compare antidepressants with each other using NNTs. The authors presented no evidence on the use of NNT and NNH in decision-making in practice.

An analysis using statistical techniques, such as indirect comparisons or mixed treatment comparisons, may facilitate a more robust comparison between two treatments.

**Implications of the review for practice and research**

**Practice:** The authors stated that SSRIs (except possibly fluoxetine) and venlafaxine are more effective than other antidepressants in the treatment of depression in older people.

**Research:** The authors stated that regulatory authorities should require future studies to present NNT and NNH data.

**Bibliographic details**


**PubMedID**

12103451

**Other publications of related interest**


**Indexing Status**

Subject indexing assigned by NLM

**MeSH**

Aged; Aged, 80 and over; Antidepressive Agents /therapeutic use; Depressive Disorder /drug therapy; Evidence-Based Medicine; Humans

**AccessionNumber**

12002008866

**Date bibliographic record published**

31/03/2005

**Date abstract record published**

31/03/2005
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.