A meta-analysis of the treatment of panic disorder with or without agoraphobia: a comparison of psychopharmacological, cognitive-behavioral, and combination treatments

van Balkom A J, Bakker A, Spinhoven P, Blaauw B M, Smeenk S, Ruesink B

Authors' objectives
To compare short-term efficacy of benzodiazepines, antidepressants, psychological panic management, exposure in vivo, and combination treatments in panic disorder with or without agoraphobia.

Searching
Excerpta Medica, PsycINFO and Index Medicus were searched for studies published between 1964 and 1995 using the keywords 'panic disorder', 'agoraphobia' and 'treatment' or 'therapy'. In addition, the references of retrieved articles were searched. A search for unpublished literature was not undertaken.

Study selection

Study designs of evaluations included in the review
'Therapy outcome studies' published between 1964 and 1995 were included. The design of the studies was unclear but it appears that a variety of designs were eligible for inclusion. The majority of the included studies did not have a control condition. Case reports, double publications, reviews, and studies that did not provide sufficient information to calculate effect sizes were excluded.

Specific interventions included in the review
High-potency benzodiazepines, antidepressants, psychological panic management, exposure in vivo, and combination treatments. The combination treatments comprised the following: pill placebo and exposure in vivo; antidepressants and exposure in vivo; and psychological panic management and exposure. Studies concerned with other psychotropic compounds or psychotherapeutic methods were excluded due to a lack of evidence.

Participants included in the review
Patients where the main diagnosis was for panic disorder with or without agoraphobia. Diagnostic criteria of the American Psychiatric Association were used in 27% (DSM-III-R criteria) and 43% (DSM-III criteria) of the treatment conditions. In the remaining 30%, patients were simply diagnosed as agoraphobic.

Outcomes assessed in the review
Panic, agoraphobia, depression, and general anxiety. A variety of assessment scales appear to have been used to assess these outcomes.

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

Assessment of study quality
Fourteen items were used to assess the quality of the studies based on an assessment of: study population characteristics, control of the interventions evaluated, relevance of the assessments, presentation of the data, and the statistical analyses performed. Four reviewers assessed the studies by filling in coding forms. Each quality criterion was rated on a 5-point Likert-type scale. The end point of these scales were defined with higher scores reflecting better quality. The scores were summed to give an overall quality score (range: 14 to 70).

Data extraction
Two reviewers independently extracted data from the studies. Problems concerning the extraction process were
discussed weekly. For each study, the two data extraction forms were compared and any discrepancies were resolved by referring to the original data or by consensus after discussion. A single final extraction form was prepared for each article.

**Methods of synthesis**

How were the studies combined?

Effect sizes (Cohen's d) were calculated within each treatment condition by comparing the mean pre-test and post-test scores for completer data; intention to treat data would have been used, but it were incomplete in 92% of the studies. A mean d was calculated for each outcome measure by averaging the effect sizes.

Differences between the treatment conditions on demographic and psychiatric status variables, the total quality score, and the magnitude of the effect sizes, were analysed with the Kruskal-Wallis one-way analysis of variance by ranks. Where significant differences in the rank distribution occurred, the Mann-Whitney test was employed. Regression analysis was used to indicate whether the differences between the conditions found in the main analyses were still present after accounting for the variance explained by the demographic, clinical status, and quality variables.

How were differences between studies investigated?

Formal statistical tests were performed to test for heterogeneity, but the methods used were not stated.

**Results of the review**

There were 106 studies pertaining to 222 treatment conditions. These involved 5,011 patients at pre-test and 4,016 at post-test.

Antidepressants, psychological panic management, high-potency benzodiazepines, and antidepressants combined with exposure in vivo, were superior to control treatments for panic attacks. Exposure in vivo alone was not effective for panic attacks. All seven treatments were superior to control conditions for agoraphobic avoidance. No difference in panic attack rates were found when comparing the treatments. For agoraphobic avoidance, the combination of antidepressants with exposure in vivo was superior to the other conditions.

**Authors' conclusions**

The combination of antidepressants with exposure in vivo was superior to all the other therapies for the short-term treatment of patients with panic disorder with agoraphobic avoidance. This should be the first treatment of choice.

**CRD commentary**

Overall, this was a well-reported and adequately conducted review. A comprehensive electronic search was undertaken, but the handsearch would have benefited from being extended to include specific key journals and contact with experts in the field. The inclusion and exclusion criteria are stated, but it was unclear what study designs were included from the description given, which refers to 'therapy outcome studies'. Considering the variety of study designs included in the review, the use of a quality assessment was particularly important. Viewing the quality assessment scores suggests that the studies were at best of average quality.

The use of the 'fail-safe N' to determine the number of unpublished (or undetected) no-effect studies needed to overturn the analysis is a useful method of assessing the strength of the review's conclusions. However, this analysis is only valid if the studies are combined appropriately. Poor internal validity (quality), the likely inclusion of several study designs, and the fact that most studies did not have a direct control comparison (i.e. the analysis was based on indirect comparisons that could introduce a significant degree of bias) suggest that the conclusions of the review should be viewed with some caution, despite the 'fail-safe N' analysis.

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

The authors did not state any clear implications for practice and further research.
Bibliographic details

PubMedID
9284865

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Adult; Agoraphobia /epidemiology /psychology /therapy; Antidepressive Agents /therapeutic use; Benzodiazepines /therapeutic use; Cognitive Therapy; Combined Modality Therapy; Female; Humans; Male; Panic Disorder /epidemiology /psychology /therapy; Treatment Outcome

AccessionNumber
11997001064

Date bibliographic record published
30/06/1999

Date abstract record published
30/06/1999

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.