Large-scale observational studies of hypericum extracts in patients with depressive disorders: a systematic review

Linde K, Knuppel L

CRD summary
This review used observational studies to assess the effects of using hypericum extracts for depressive disorders. The authors concluded that hypericum appeared to be effective and well-tolerated for the routine treatment of patients with mild to moderate depression. Given the limitations of evidence from the identified non-randomised studies, the authors' cautious conclusion appears appropriate.

Authors' objectives
To use observational studies to assess the effects of hypericum extracts in the treatment of depressive disorders.

Searching
MEDLINE (1966 to November 2002) and PubMed (to March 2003) were searched using the reported keywords. In addition, manufacturers were contacted and proceedings of herbal medicine congresses were screened.

Study selection
Study designs of evaluations included in the review
Observational non-experimental studies with at least 100 patients were eligible for inclusion. Randomised controlled trials were excluded.

Specific interventions included in the review
Studies of hypericum monotherapy in primary care settings were eligible for inclusion. Studies using hypericum in combination with other plant extracts were excluded. The included studies used 12 different products based on the same extracts, and the daily extract dose ranged from about 360 to 1,200 mg (the details were reported). Most of the studies lasted 4 to 6 weeks (short term); two studies lasted for one year (long term).

Participants included in the review
Studies of patients suffering primarily from depressive disorders were eligible for inclusion. Studies that included patients with other psychiatric disorders were included if the majority of the patients had depression. Studies of patients with other medical conditions were excluded. The included studies were undertaken predominantly in patients with mild to moderately severe depression; some studies also included a number of patients with severe depression or other psychiatric syndromes. One study was in children less than 12 years old.

Outcomes assessed in the review
Studies that assessed clinical outcomes such as response, change on depression rating scales and adverse effects were eligible for inclusion. The review assessed the percentage of physician-rated and patient-rated responders (defined as good/very good efficacy or good/very good improvement, or as defined using a rating scale) in the short and long term, as well as adverse effects (defined as any undesired event), side-effects (adverse event with possible causal link to hypericum) and drop-outs due to side-effects. The results were reported for physician-rated assessments using the Hamilton rating scale for depression, and for patient-rated assessments using the Depression scale von Zerssen.

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

Assessment of study quality
The studies were assessed on the basis of: description of recruitment methods and reporting of inclusion criteria; the likelihood of the study sample representing patients in routine practice; diagnosis made using a common classification system; adequate reporting of patient characteristics; observation for at least 6 weeks; description of cointerventions and/or co-morbidity; use of validated patient-rated and physician-rated instruments to assess the severity of depression; before-and-after data reported as point estimates and measures of variability; description of drop-outs and withdrawals; critical discussion of study methods and results; and adequate analysis of prognostic factors with adjustment for confounders. One reviewer assessed validity.

Data extraction
Two reviewers extracted data from all of the studies, using a data extraction form, and one of these reviewers cross-checked the data. Data were extracted on the proportion of patients classified as responders, the number of patients with adverse events, the total number of adverse events, the number of patients reporting side-effects, the total number of side-effects and the number of patients dropping out due to side-effects.

Methods of synthesis
How were the studies combined?
The studies were combined in a narrative and the range of percentage event rates was reported.

How were differences between studies investigated?
Differences between the studies were discussed in the narrative synthesis, with respect to study quality criteria in particular.

Results of the review
Sixteen observational studies (n=34,804) were included.

In terms of study quality, the major methodological flaw was the lack of reporting of relevant information. Seven studies used a validated depression rating scale. Most of the studies reported drop-outs and withdrawals. Ten studies assessed prognostic factors but none adjusted for confounders. Most of the studies did not describe the methods used to record adverse events and side-effects.

The studies showed physician-rated response rates ranging from 65 to 100% in the short term and 60 to 69% (based on two studies) in the long term.

The studies showed patient-rated response rates ranging from 63 to 98% in the short term. In the one study assessing long-term outcomes, 85% rated efficacy as good or excellent.

Drop-out rates ranged from 1.5 to 17.1% in the short term; the rate was 19.8% in the one study assessing long-term outcomes.

Drop-outs due to side-effects were rare, ranging from 0 to 2.8% in short-term studies and from 3.4 to 5.7% in the two long-term studies. The proportion of patients reporting side-effects was also low, ranging from 0 to 5.9%.

The most common side-effects were gastrointestinal symptoms, increased sensitivity to light and skin symptoms, although several studies reported a range of nervous symptoms. None of the studies reported any serious side-effects requiring hospitalisation.

The results from studies examining prognostic factors were also reported in the review.

Authors’ conclusions
Hypericum appeared to be effective and well-tolerated for the routine treatment of patients with mild to moderate depression.
CRD commentary
The review addressed a clear question that was defined in terms of the participants, intervention, outcomes and study design; the inclusion criteria for the outcomes were broad. Only one electronic database was searched, but this was supplemented by contact with manufacturers and searches of conference papers, which reduces the possibility of publication bias. It appears that attempts were made to minimise language bias, although this was not explicitly stated. The authors highlighted the difficulty in locating relevant observational studies and acknowledged the possibility that studies were missed. Methods were used to minimise errors and bias in the extraction of data but not in the assessment of validity. However, the methods used to select studies were not described, so it is not known whether any efforts were made to reduce errors and bias. Validity was assessed using specified criteria, the results of the assessment were reported, and validity was taken into account when considering the results. Given the weaknesses of these studies and the lower reliability of evidence from non-randomised studies, the authors’ cautious conclusion appears appropriate.

Implications of the review for practice and research
Practice: The authors did not state any implications for practice.

Research: The authors stated that future research should aim to examine prognostic factors and should possibly reanalyse available studies.

Bibliographic details

PubMedID
15693723

DOI
10.1016/j.phymed.2004.02.004

Subject indexing assigned by NLM
Antidepressive Agents /therapeutic use; Depressive Disorder /drug therapy; Humans; Hypericum; Phytotherapy; Plant Extracts /therapeutic use; Psychiatric Status Rating Scales; Treatment Outcome

AccessionNumber
12005009230

Date bibliographic record published
31/07/2006

Date abstract record published
31/07/2006

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.