St John's wort for major depression (Review)

Linde K, Berner MM, Kriston L

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2008, Issue 4

http://www.thecochranelibrary.com

TABLE OF CONTENTS

HEADER . . . 1
ABSTRACT . . . 1
PLAIN LANGUAGE SUMMARY . . . 3
BACKGROUND . . . 3
OBJECTIVES . . . 4
METHODS . . . 4
RESULTS . . . 7
Figure 1. . . . 10
Figure 2. . . . 11
Figure 3. . . . 12
Figure 4. . . . 14
Figure 5. . . . 15
Figure 6. . . . 16
Figure 7. . . . 18
DISCUSSION . . . 19
AUTHORS' CONCLUSIONS . . . 20
ACKNOWLEDGEMENTS . . . 21
REFERENCES . . . 21
CHARACTERISTICS OF STUDIES . . . 27
DATA AND ANALYSES . . . 50
Analysis 1.1. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 1 Responder - grouped by precision - primary analysis. . . . 55
Analysis 1.2. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 2 Responder - according to HAMD. . . . 58
Analysis 1.3. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 3 Responder - according to CGI (Clinical Global Impression Index at least "much improved"). . . . 61
Analysis 1.4. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 4 Responder - grouped by extract. . . . 63
Analysis 1.5. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 5 Responder among studies from German-speaking countries and other studies. . . . 67
Analysis 1.6. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 6 Remission (HAMD score < 8 or < 7). . . . 71
Analysis 2.1. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 1 Mean HAMD (Hamilton Rating Scale for Depression) scores after therapy. . . . 72
Analysis 2.2. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 to 3 weeks of treatment. . . . 73
Analysis 2.3. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 3 Mean HAMD (Hamilton Rating Scale for Depression) score after 4 weeks of treatment. . . . 74
Analysis 2.4. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment. . . . 75
Analysis 2.5. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment. . . . 76
Analysis 2.6. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 6 MADRS after treatment. . . . 77
Analysis 2.7. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 7 Difference MADRS baseline - end of treatment. . . . 77
Analysis 2.8. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 8 Mean HAMD after treatment in studies from German-speaking countries and other studies. . . . 78
Analysis 2.9. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 9 Mean Depression Scale von Zerssen (D-S) after therapy/difference baseline - after therapy. . . . 80

Analysis 2.10. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 10 Various self-rating scales. 81
Analysis 2.11. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 11 Various self-rating scales in studies from German-speaking countries and other countries. 86
Analysis 3.1. Comparison 3 Safety - Hypericum mono-preparations vs. placebo, Outcome 1 Number of patients discontinuing treatment/dropping out for adverse effects - primary analysis. 88
Analysis 3.2. Comparison 3 Safety - Hypericum mono-preparations vs. placebo, Outcome 2 Number of patients dropping out. 89
Analysis 3.3. Comparison 3 Safety - Hypericum mono-preparations vs. placebo, Outcome 3 Number of patients reporting adverse effects. 90
Analysis 4.1. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 1 Responder (intent to treat) - primary analysis. 91
Analysis 4.2. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 2 Responder (per protocol). 94
Analysis 4.3. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 3 Responders according to CGI (Clinical Global Impression Index at least “much improved”). . . 97
Analysis 4.4. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 4 Responder among studies from German-speaking studies and other studies. 100
Analysis 4.5. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 5 Remission (HAMD score < 8). 103
Analysis 5.1. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 1 Mean HAMD (Hamilton Rating Scale for Depression) after therapy. 104
Analysis 5.2. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 or 3 weeks of treatment. . . 107
Analysis 5.3. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 3 Mean HAMD (Hamilton Rating Scale for Depression) scores after 4 weeks of treatment. . . . 109
Analysis 5.4. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment. . . 111
Analysis 5.5. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment. 113
Analysis 5.6. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 6 MADRS after treatment. 115
Analysis 5.7. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 7 Difference MADRS baseline - end of treatment. 116
Analysis 5.8. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 8 Mean HAMD after treatment in studies from German-speaking countries and other studies. . . 117
Analysis 5.9. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 9 Mean D-S (Depression Scale von Zerssen) scores after therapy. 119
Analysis 5.10. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 10 Various self-rating scales. 121
Analysis 5.11. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 11 Various self-rating scales in studies from German-speaking countries and other countries. . . . 125
Analysis 6.1. Comparison 6 Safety - Hypericum mono-preparations vs. standard antidepressants, Outcome 1 Number of patients discontinuing treatment/dropping out due to adverse/side effects. 127
Analysis 6.2. Comparison 6 Safety - Hypericum mono-preparations vs. standard antidepressants, Outcome 2 Number of patients dropping out. 129
Analysis 6.3. Comparison 6 Safety - Hypericum mono-preparations vs. standard antidepressants, Outcome 3 Number of patients reporting adverse effects. 132
APPENDICES 134
WHAT’S NEW 141
HISTORY 142
CONTRIBUTIONS OF AUTHORS 142
DECLARATIONS OF INTEREST 142

SOURCES OF SUPPORT 142
INDEX TERMS 142

[Intervention review]

St John's wort for major depression

Klaus Linde[1], Michael M Berner[2], Levente Kriston[2]

[1]Centre for Complementary Medicine Research, Department of Internal Medicine II, Technische Universitaet Muenchen, Munich, Germany. [2]Department of Psychiatry and Psychotherapy, University Medical Center Freiburg, Freiburg im Breisgau, Germany

Contact address: Klaus Linde, Centre for Complementary Medicine Research, Department of Internal Medicine II, Technische Universitaet Muenchen, Wolfgangstr. 8, Munich, 81667, Germany. Klaus.Linde@lrz.tu-muenchen.de. (Editorial group: Cochrane Depression, Anxiety and Neurosis Group.)

Cochrane Database of Systematic Reviews, Issue 4, 2008 (Status in this issue: *Edited, conclusions changed*)

DOI: 10.1002/14651858.CD000448.pub3
This version first published online: 8 October 2008 in Issue 4, 2008. Re-published online with edits: 8 October 2008 in Issue 4, 2008.
Last assessed as up-to-date: 7 August 2008. (Dates and statuses?)

This record should be cited as: Linde K, Berner MM, Kriston L. St John's wort for major depression. *Cochrane Database of Systematic Reviews* 2008, Issue 4. Art. No.: CD000448. DOI: 10.1002/14651858.CD000448.pub3.

ABSTRACT

Background

In some countries extracts of the plant *Hypericum perforatum* L. (popularly called St. John's wort) are widely used for treating patients with depressive symptoms.

Objectives

To investigate whether extracts of hypericum are more effective than placebo and as effective as standard antidepressants in the treatment of major depression; and whether they have fewer adverse effects than standard antidepressant drugs.

Search strategy

Trials were searched in computerised databases, by checking bibliographies of relevant articles, and by contacting manufacturers and researchers.

Selection criteria

Trials were included if they: (1) were randomised and double-blind; (2) included patients with major depression; (3) compared extracts of St. John's wort with placebo or standard antidepressants; (4) included clinical outcomes assessing depressive symptoms.

Data collection and analysis

At least two independent reviewers extracted information from study reports. The main outcome measure for assessing effectiveness was the responder rate ratio (the relative risk of having a response to treatment). The main outcome measure for adverse effects was the number of patients dropping out due to adverse effects.

Main results

A total of 29 trials (5489 patients) including 18 comparisons with placebo and 17 comparisons with synthetic standard antidepressants met the inclusion criteria. Results of placebo-controlled trials showed marked heterogeneity. In nine larger trials the combined response rate ratio (RR) for hypericum extracts compared with placebo was 1.28 (95% confidence interval (CI), 1.10 to 1.49) and from nine smaller trials was 1.87 (95% CI, 1.22 to 2.87). Results of trials comparing hypericum extracts and standard antidepressants were statistically homogeneous. Compared with tri- or tetracyclic antidepressants and selective serotonin reuptake inhibitors (SSRIs),

respectively, RRs were 1.02 (95% CI, 0.90 to 1.15; 5 trials) and 1.00 (95% CI, 0.90 to 1.11; 12 trials). Both in placebo-controlled trials and in comparisons with standard antidepressants, trials from German-speaking countries reported findings more favourable to hypericum. Patients given hypericum extracts dropped out of trials due to adverse effects less frequently than those given older antidepressants (odds ratio (OR) 0.24; 95% CI, 0.13 to 0.46) or SSRIs (OR 0.53, 95% CI, 0.34-0.83).

Authors' conclusions

The available evidence suggests that the hypericum extracts tested in the included trials a) are superior to placebo in patients with major depression; b) are similarly effective as standard antidepressants; c) and have fewer side effects than standard antidepressants. The association of country of origin and precision with effects sizes complicates the interpretation.

PLAIN LANGUAGE SUMMARY

St. John's wort for treating depression.

Depression is characterised by depressed mood and/or loss of interest or pleasure in nearly all activities and a variety of other symptoms for periods longer than two weeks. Extracts of St. John's wort (botanical name *Hypericum perforatum L.*) are prescribed widely for the treatment of depression.

We have reviewed 29 studies in 5489 patients with depression that compared treatment with extracts of St. John's wort for 4 to 12 weeks with placebo treatment or standard antidepressants. The studies came from a variety of countries, tested several different St. John's wort extracts, and mostly included patients suffering from mild to moderately severe symptoms. Overall, the St. John's wort extracts tested in the trials were superior to placebo, similarly effective as standard antidepressants, and had fewer side effects than standard antidepressants. However, findings were more favourable to St. John's wort extracts in studies form German-speaking countries where these products have a long tradition and are often prescribed by physicians, while in studies from other countries St. John's wort extracts seemed less effective. This differences could be due to the inclusion of patients with slightly different types of depression, but it cannot be ruled out that some smaller studies from German-speaking countries were flawed and reported overoptimistic results.

Patients suffering from depressive symptoms who wish to use a St. John's wort product should consult a health professional. Using a St. John's wort extract might be justified, but important issues should be taken into account: St. John's wort products available on the market vary to a great extent. The results of this review apply only to the preparations tested in the studies included, and possibly to extracts with similar characteristics. Side effects of St. John's wort extracts are usually minor and uncommon. However, the effects of other drugs might be significantly compromised.

BACKGROUND

Description of the condition

Depressive disorders are characterised by depressed mood and/or loss of interest or pleasure in nearly all activities in the presence of other symptoms such as loss of appetite, fatigue and lack of energy, sleep disturbance, restlessness or irritability, feelings of worthlessness or inappropriate guilt, difficulty in thinking, concentrating or making decisions and thoughts of death or suicide or attempts at suicide (Candy 2008). Depressive disorders are the largest source of non-fatal disease burden in the world, accounting for 12% of years lived with disability (Ustun 2004). There are two major classification systems to diagnose depressive disorders, the Diagnostic and Statistical Manual of Mental Disorders (DSM; current version DSM-IV-TR) and the International Statistical Classification of Diseases and Related Health Problems (ICD; current version ICD-10). DSM-IV defined depressive diagnoses include recurrent or persistent major depression and minor depression. ICD-10 diagnoses include recurrent or persistent depression with mild, moderate or severe episodes. According to the DSM-IV diagnostic classification, either depressed mood or a loss of interest or pleasure in daily activities consistently for at least a two week period has to be present to diagnose a major depressive disorder. The ICD-10 system uses the term depressive episode instead of major depressive disorder, but lists similar criteria.

Description of the intervention

Major depressive episodes are most commonly treated with antidepressant medication. Current first line dugs are selective serotonin reuptake inhibitors (SSRI) or tricyclic and related antidepressants (http://guidance.nice.org.uk/CG23). However, the size of effects over placebo in clinical trials has been modest (Turner 2008; Kirsch 2008), and although SSRIs are better tolerated than older antidepressants, side effects still occur in a relevant proportion of patients.

Extracts of the plant *Hypericum perforatum* L. (St. John's wort), a member of the Hypericaceae family, have been used in folk medicine for a long time for a range of indications including depressive disorders. Extracts of St. John's wort are licensed and widely used in Germany for the treatment of depressive, anxiety and sleep disorders. In recent years, hypericum extracts have also become increasingly popular in other countries.

How the intervention might work

The exact mechanism of action of the antidepressant effects of hypericum extracts is still unclear. Hypericum extracts contain at least seven constituents or groups of components that may contribute to its pharmacological effects (Nahrstedt 1997). These include naphthodianthrons (e.g., hypericins), flavonoids (e.g., quercetin), biflavonoids (e.g., biapigenin), xanthons, and phloroglucinol derivatives (e.g., hyperforin). Hypericum extracts have been shown to be active in a number of standard animal models that are used to indicate antidepressant effects (Wheatley 1998; Caccia 2005; Wurglics 2006). While some isolated substances, as for example hyperforin, have been shown to have antidepressant activity, the total extract seems to be more effective (Reichling 2003).

Why it is important to do this review

Hypericum extracts have been tested in a number of clinical trials since the 1980s. The first two versions of this review and other systematic reviews published between 1995 and 2000 concluded that hypericum extracts are more effective than placebo and are comparable to older antidepressants in the treatment of mild to moderate depressive disorders (Ernst 1995; Linde 1996; Linde 1998; Kim 1999; Gaster 2000; Williams 2000). Several trials included in these reviews were criticised because they included patients with few and/or mild symptoms who did not meet criteria for major depression, were conducted by primary care physicians who were not experienced in depression research, and/or used low doses of comparator drugs (Shelton 2001). In the 2005 update of our review (Linde 2005a; Linde 2005b) several new well-designed placebo-controlled trials were included, some of which had negative findings (Shelton 2001; HDTSG 2002) and which had spurred renewed debate about the efficacy of hypericum extracts. We systematically investigated possible reasons for the contradictory findings. We found that larger, more precise studies yielded less positive results, suggesting that small studies with a higher risk of bias might overestimate the effects of hypericum extracts over placebo. The analyses also showed that effects over placebo were less pronounced in studies restricted to patients with major depression. Finally, we had the impression that studies originating from German-speaking countries (Germany, Austria, and Switzerland) had more positive results than studies originating from other countries independently from precision and formal diagnosis, although multiple regression analyses did not identify this as an independent predictor.

Since we completed the search for our 2005 update, again, several new well-designed trials restricted to patients with major depression have been published. To sharpen the focus of this review, to reduce clinical heterogeneity, and to reflect the fact that almost all new high-quality trials of hypericum extracts are restricted to patients with major depression, we decided to limit the review now to this group of patients.

OBJECTIVES

This updated review aimed to investigate whether extracts of hypericum:

- are more effective than placebo and
- as effective as standard antidepressant drugs, and
- whether they have less adverse effects compared to standard antidepressant drugs

in the treatment of major depression in adults.

In addition, we investigated possible reasons for varying results across studies, with a focus on precision of the studies, baseline severity of depression, and country of origin.

METHODS

Criteria for considering studies for this review

Types of studies

To be included trials had to be double-blind and randomised.

Types of participants

Patients had to suffer from major depression (meeting DSM-IV or ICD-10 criteria). Trials in children (< 16 years) were not eligible. In previous versions of this review (Linde 1998; Linde 2005a) trials not restricted to patients with major depression had been included.

Types of interventions

Experimental intervention

Preparations of hypericum (St. John's wort). Trials investigating combinations of hypericum with other herbs were excluded.

Control intervention

Placebo or synthetic antidepressants (tricyclic and related antidepressants, selective serotonine reuptake inhibitors, serotonine-noradrenaline reuptake inhibitors). Trials using clearly inadequate synthetic antidepressants (e.g., benzodiatepines) or a dosage clearly below the lower thresholds recommended in current guidelines (Härter 2003, ICSI 2007) were excluded.

Experimental and control treatments had to be given for at least four weeks.

Main comparisons

The following comparisons were performed:

1. hypericum extracts vs. placebo
2. hypericum extracts vs. standard antidepressants

Types of outcome measures

Primary outcome

To be included, trials had to measure clinical outcomes such as depression scales or symptoms. Trials that measured physiological parameters only were excluded. The primary outcomes of interest were

1. Effectiveness: treatment response
2. Safety: the proportion of patients who dropped out due to adverse effects

Secondary outcomes

1. Effectiveness: remission, depression scales such as the Hamilton Depression Scale (HAMD), the Clinical Global Impression

Index (CGI), the Montgomery-Asberg Depression Rating Scale (MADRS), patient-rated depression scales
2. Safety: total proportion of drop-outs, proportion of patients reporting adverse effects

Search methods for identification of studies

For the first version of the review we searched for published and unpublished eligible trials in the following ways:
1. Electronic searches
a) Clinical Trials Register of the Cochrane Collaboration Depression, Anxiety & Neurosis Group (CCDANTR)
b) database of the Cochrane Field for Complementary Medicine
c) full text searches in Medline SilverPlatter CD-ROM from 1983 onwards and Embase 1989 onwards using the terms 'St. John's wort', 'Johanniskraut' (German for St John's wort), 'hyperic*')
d) full text searches in Psychlit and Psychindex 1987 - 1997 CD-ROM
e) searches in the private database Phytodok, Munich.
2. Searching other resources
a) Checking bibliographies of obtained articles
b) Contacting pharmaceutical companies and authors.
There were no language restrictions.
For the updated version of the review, we searched for published and unpublished eligible trials in the following ways:
1. Electronic searches
For the update, regular electronic searches were performed in CCDANTR (last search July 2007) and PubMed (screening all hits for text word "hypericum", last search July 8, 2008).
2. Searching other resources
We screened bibliographies of published articles, and repeatedly contacted experts, researchers, and manufacturers inquiring for new trials. One reviewer (KL) initially screened reference lists to identify controlled studies on hypericum preparations in humans. All possibly relevant studies or publications were then checked formally for eligibility.

Data collection and analysis

Selection of studies

Two reviewers independently decided on eligibility for the revised inclusion criteria. Disagreements were resolved through discussion. Due to reading errors, disagreements occurred for two trials in which not all patients had major depression and which were excluded after discussion (Vorbach 1994, Winkel 2000). In two trials both reviewers had problems with assessing eligibility: For one small, older trial (Lehrl 1993) the publication did not state that inclusion was limited to patients with major depression, but we had a statement of the sponsor obtained for our 1998 update that all patients met the criteria. As this information could not be verified for this update, we decided to exclude the trial. A Chinese trial (Gu 2001) referred to a Chinese classification. As this classification is not completely comparable to ICD-10 and DSM-IV, we decided to also exclude this trial.

Data extraction and management

Primary study characteristics and results were extracted by at least two independent reviewers using a pretested form. In particular, we extracted diagnoses and main inclusion criteria, age, gender, duration of episodes, baseline depression scores, country of origin, number and type of study centers, numbers of patients who were randomised and analysed and who completed protocols, the number and reasons for drop-outs and withdrawals, numbers of patients reporting adverse effects, and the number and type of adverse effects that were reported.
We assessed numbers of patients who were classified as responders based on score improvements on the Hamilton Rating Scale for Depression (HAMD), the Clinical Global Impression Index (CGI; subscale global improvement rating as at least "much improved"), or any other clinical response measurement. Missing or additional information was sought from authors/sponsors.
Most trials measured clinical outcomes with the Hamilton Depression Scale (HAMD) and the Clinical Global Impression Index (CGI). The HAMD is an observer-rated scale that focuses mainly on somatic symptoms of depression (Hamilton 1960). The original version includes 21 items, but a version with 17 items is more commonly used in clinical trials. Most studies using the HAMD report the number of 'treatment responders' (patients achieving a score less than 10 and/or less than 50% of the baseline score). When available, we extracted means and standard deviations before, during and after treatment as well as the number of 'responders'. The CGI (CGI 1970) is an observer rated instrument with three items (severity of illness, global improvement, and an efficacy index). We extracted the number of patients rated as 'much improved' or 'very much improved' for global improvement. As recently the Montgomery-Asberg Depression Rating Scale (MADRS (Montgomery 1979)) and the remission criterion for the HAMD (usually a score of less than 8 at the end of treatment) have been gaining importance as outcome criteria, we also checked all trials for the reporting of these measures. As the DS (Depression Scale von Zerssen (von Zerssen 1996)) was the most often used patient-rated instrument in the included trials, we extracted post-treatment data for this scale, if available. For additional post-hoc analyses, one reviewer (KL) also extracted data for other self-rating instruments.

Assessment of risk of bias in included studies

The main part of the update process of this review was completed before the new risk of bias tool of the Cochrane Collaboration (Higgins 2008) was available. The methodological quality of each trial was assessed by at least two independent reviewers using scales developed by A. Jadad et al. (Jadad 1996) and by one of the

reviewers (KL). The results of the quality scoring are displayed in the table of included studies.

The Jadad scale has three items adding up to a maximum score of five points. 0, 1 or 2 points can be given for randomisation (explicit statement that allocation was randomised and description of an adequate generation of the random sequence), 0, 1 or 2 points for double-blinding (explicit statement that patients and evaluators were blinded and that treatments were indistinguishable), 0 or 1 point for description of drop-outs and withdrawals (numbers and reasons for all compared groups separately). The display in the table of included studies is as follows (examples): 2-2-1 (full score in every item), 1-0-0 (only statement on randomisation).

The second quality scale, the "Internal Validity Scale" (IV), which has been used in other reviews on complementary medicine (Linde 1996b, Linde 1997) has six items with possible scores of 0, 0.5 or 1 point for each. Items 1 through 6 refer to statement of random allocation, adequacy of randomization concealment, baseline comparability, blinding of patients, blinding of evaluators, and likelihood of selection bias after allocation, respectively. Results are displayed by item in the table of included studies (e.g., 1-1-1-0.5-1-1 represents a full score, with the exception of blinding of patients which was stated but treatment and placebo might have been distinguishable).

The assessments in the Jadad and IV scores are solely based on the information provided in the publication (as additional information could not be gathered for all studies). In the table 'Characteristics of included studies', however, additional information provided from authors or sponsors was included. This table also contains information on allocation concealment and attrition.

Measures of treatment effect

Our primary outcome measure, to assess the effectiveness of St John's wort versus placebo and versus other antidepressants, was the proportion of responders (according to the Hamilton Depression Scale (first preference) or other responder measures (second preference)) at the end of treatment, or in case of treatment phases longer than 6 weeks, at the time point defined for primary outcome measurement by the study investigators.

Secondary outcome measures were: proportion of responders according to HAMD, proportion of responders according to CGI, mean HAMD after treatment (or, if this was not available, difference after treatment - baseline), at 2, 4, 6 to 8 weeks, and mean DS score after treatment (or, if this was not available, difference after treatment - baseline).

The main outcome measure for the safety analysis was the proportion of patients who dropped out due to adverse effects. Secondary measures were the total proportion of drop-outs and the proportion of patients reporting adverse effects.

Dichotomous outcomes

We used responder rate ratios (= relative risks = proportion of responders in the treatment group/proportion of responders in the control group) and their 95% confidence intervals for the analysis of treatment response. Responder rate ratios greater than 1 indicate better response in the hypericum group.

Due to highly variable frequency of side or adverse effects reported, odds ratios instead of rate ratios were calculated in the safety analyses. Odds ratios less than 1 indicate that fewer events occurred in the hypericum group.

Continuous outcomes

For HAMD and DS scores we calculated mean differences (also termed weighted mean differences). Negative mean differences indicate better response in the hypericum group.

Unit of analysis issues

Two trials with more than one hypericum group were included in the analyses as follows: Laakmann 1998 included an extract available on the market and an additional experimental extract with low hyperforin content which was never on the market and only used for control reasons. We did not include the data from the group receiving the experimental low-hyperforin extract in the analyses. Kasper 2006 et al tested two dosages (600 and 1200 mg) of an available product. We pooled the data from these two groups to prevent that the control group of this trial would have been included in the analyses twice.

Dealing with missing data

Dichtomous outcomes

Responder proportions were calculated according to the intention to treat principle, counting drop-outs as non-responders. For the comparison hypericum extracts vs. standard antidepressants responder proportions were also calculated on a per protocol basis (as this is considered more appropriate to assess the equivalence of two treatments).

Continuous outcomes

If means and standard deviations from intent to treat analysis with missing values replaced were available, we preferably used these data. In other cases we used analysis based on available data.

Obtaining missing data

If the number of patients responding to treatment and means and/or standard deviations of HAMD scores after completion were not reported, we always tried to contact first or corresponding authors and/or sponsors to obtain these data. In general, we also tried to obtain other missing details on methods and secondary outcomes from authors or sponsors, but the extent to which we were doing this depended on the cooperation of authors/sponsors and the amount of missing information in the publications. We did not impute or recalculate missing standard deviations as these were unavailable only for a few secondary outcomes in a minority of trials.

We tried to contact authors and/or sponsors of 27 of the 29 included trials; for two trials (Kalb 2001; Laakmann 1998) this was considered unnecessary. We did not receive responses for five trials (Behnke 2002; Brenner 2000; Fava 2005; Harrer 1999; Moreno 2005). Very limited additional information was available or needed for three studies (Bjerkenstedt 2005; Volz 2000; Woelk

2000). We obtained relevant additional information to a variable extent from authors, sponsors, or both for the remaining 19 trials.

Data synthesis

The following comparisons were performed:
1. hypericum extracts vs. placebo: a) for dichotomous outcomes (response rate ratios); b) for continuous outcomes; c) for drop-outs and adverse effects
2. hypericum extracts vs. standard antidepressants: a) for dichotomous outcomes; b) for continuous outcomes; c) for drop-outs and adverse effects
All main analyses were performed using RevMan 5.
Due to the clinical diversity of the studied populations, the hypericum extracts and the comparison drugs used, we considered that the included studies did not estimate a common underlying effect, but rather that each individual study estimated its single and unique underlying effect. Thus, the application of random effects model in all analyses seemed to be appropriate.
The primary analysis for the comparison of response rate ratios (= relative risks) under treatment with hypericum extracts or placebo was a random-effects intention to treat meta-analysis stratified by study precision (above or below median of variance of treatment effect).

Assessment of heterogeneity

Heterogeneity of trials' results was tested with the Chi-squared test, and the I-squared statistic was calculated to give an estimate of the degree of heterogeneity. I-squared values over 50% indicate considerable heterogeneity (Higgins 2003).

Investigation of heterogeneity and subgroup analyses

Predefined subgroup analyses were performed (a) including only trials with response operationalised with the HAMD score; (b) including only trials with response operationalized with the CGI; (c) for the type of extract investigated; and (d) comparing trials originating from German-speaking countries and from other countries.
Weighted mean differences for HAMD scores were calculated after therapy, at 2 to 3, 4, 6 to 8 weeks, and for differences compared to baseline values. For DS scores we calculated after therapy values, for MADRS score after therapy values and differences compared to baseline. As only relatively few studies used the DS, we performed an additional post-hoc random effects analysis calculating standardised mean differences for any available patient-rating scale (preferably end of treatment values, but if these were not available also differences from baseline) to investigate whether findings from physician-rated instruments could be broadly reproduced.
The primary analysis for the comparison of responder rate ratios under treatment with hypericum extracts or standard antidepressants was a random effects intent to treat meta-analysis stratified for type of synthetic antidepressant (selective serotonine inhibitors or older antidepressants). Predefined subgroup analyses were performed (a) using per protocol data; (b) stratified for country (German-speaking Europe versus other countries); (c) including only trials with response operationalised with the HAMD score; and (d) including only trials with response operationalised with the CGI.
Additional meta-regression analyses were performed to investigate the influence of country of origin (German-speaking versus not German speaking), precision and HAMD baseline values on study findings (responder ratio and mean difference of HAMD scores after treatment) both in placebo and standard antidepressants comparisons. According to current recommendations of experts (Thompson 1999, Lipsey 2000), random effects meta-regression analyses were carried out using the restricted information maximum likelihood (REML) method. A main advantage of this approach is that it accounts for residual between-trial heterogeneity. Both univariable and multiple regression models were fitted. We calculated the proportion of explained heterogeneity variance by dividing the heterogeneity explained by the independent variable(s) through the total heterogeneity variance present in random-effects meta-analysis. When referring to a whole model, this coefficient was termed R^2. When referring to the contribution of single covariates the coefficient was termed β^2. In univariable meta-regression analyses these coefficients are mathematically equal. In multiple meta-regression analyses, sum of β^2 values for all covariates may be slightly different from R^2. For all meta-regression analyses the Statistical Package for the Social Sciences (SPSS; Chicago, Illinois) v13.0 software using additional macros by Wilson (Wilson 2002; Lipsey 2000) was used.

Assessment of reporting biases

Visual analysis of funnel plots was performed to identify possible publication bias (Sterne 2001). Furthermore, the asymmetry coefficient was calculated for formal examination of publication bias (Egger 1997).

RESULTS

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

Results of the search

A total of 79 possibly relevant studies were identified and checked formally for eligibility.

Included studies

Twenty nine trials including a total of 5489 (range 30 to 388) patients met inclusion criteria (see Characteristics of included studies).
Eighteen trials had a placebo-control group (Bjerkenstedt 2005; Bracher 2001; Fava 2005; Gastpar 2006; HDTSG 2002; Hänsgen 1996; Kalb 2001; Kasper 2006; Laakmann 1998; Lecrubier 2002; Montgomery 2000; Moreno 2005; Philipp 1999; Schrader 1998; Shelton 2001; Uebelhack 2004; Volz 2000; Witte 1995), and 17

trials compared hypericum with standard antidepressants (Behnke 2002; Bjerkenstedt 2005; Brenner 2000; Fava 2005; Gastpar 2005; Gastpar 2006; Harrer 1993; Harrer 1999; HDTSG 2002; Moreno 2005 ; Philipp 1999 ; Schrader 2000 ; Szegedi 2005 ; van Gurp 2002 ; Vorbach 1997 ; Wheatley 1997 ; Woelk 2000). Six trials had both a placebo and a standard antidepressant control group (Bjerkenstedt 2005 ; Fava 2005 ; Gastpar 2006 ; HDTSG 2002 ; Moreno 2005; Philipp 1999). Eight trials are newly included since the last update (Bracher 2001; Fava 2005; Gastpar 2005; Gastpar 2006 ; Kasper 2006 ; Moreno 2005 ; Szegedi 2005 ; Uebelhack 2004) and one trial which had been included based on an abstract reference only is now included fully (Bjerkenstedt 2005). These eight new trials included a total of 1947 (range 72 to 388) patients. Details on patients, methods, interventions, outcomes, and results of all included studies are described in the table of included studies.

Types of participants

The severity of depression was described as mild to moderate in 19 trials, and as moderate to severe in 9 trials (one trial did not classify severity). Eighteen trials were from German-speaking countries, four from the US, two from the UK, and one each from Brazil, Canada, Denmark, France and Sweden. Patients were recruited in private practices in all trials from German-language countries, in the trials from Sweden (Bjerkenstedt 2005) and Canada (van Gurp 2002), and in one of the trials from the UK (Wheatley 1997). The second trial from the UK (Montgomery 2000) and the trial from France (Lecrubier 2002) were performed both in private practices and psychiatric outpatient departments. Three trials from the US (Shelton 2001; HDTSG 2002; Fava 2005) and the Brazilian trial (Moreno 2005) were performed in academic and/or community psychiatry research clinics. Two trials from the US and Denmark (Brenner 2000; Behnke 2002) did not report on the setting.

Types of intervention

A variety of hypericum preparations were studied in the trials. The range of daily extract doses varied between 240 and 1800 mg, but in most trials 500 to 1200 mg were used. The standard antidepressants used as active comparators were fluoxetine (6 trials, dosage 20 to 40 mg), sertraline (4 trials, 50 to 100 mg), imipramine (in 3 trials, dosage 100 to 150 mg), citalopram (1 trial, 20 mg), paroxetine (1 trial, 20 to 40 mg), maprotiline (1 trial, 75 mg), and amitriptyline (1 trial, 75 mg). The comparator dosage of maprotiline and amitriptyline were slightly below of those recommended in current guidelines (Härter 2003, ICSI 2007) and in most other studies at the minimum of recommended dosages. The treatment periods lasted 4 (1 trial), 6 (19), 7 (1), 8 (5) or 12 weeks (4 trials). Four trials included some long-term follow-up or continuation treatment after the main trial phase (Brenner 2000; Gastpar 2005; Shelton 2001; Szegedi 2005).

Types of outcome

The most frequently used instrument used for outcome measurement was the Hamilton Rating Scale for Depression (used in all trials). A variety of other ratings scales and instruments were used in addition.

Excluded studies

Fifty trials (see Characteristics of excluded studies) did not meet inclusion criteria: eight trials were not limited to patients with depression (Albertini 1986; Bendre 1980; Dittmer 1992; Hottenrott 1997 ; Maisenbacher 1995 ; Panijel 1985 ; Sindrup 2000 ; Volz 2002), four trials were on prevention or treatment of depressive symptoms in patients suffering primarily from other diseases (Häring 1996 ; Li 2005 ; Mo 2004 ; Werth 1989), two measured physiological parameters only (as EEG) in depressed patients (Czekalla 1997; Kugler 1990b), five did not include a placebo or standard drug comparison group (Bernhadt 1993; Lenoir 1999; Martinez 1993 ; Spielberger 1985 ; Zeller 2000), eight involved healthy volunteers (Brockmöller 1997 ; Herberg 1992 ; Johnson 1992 ; Johnson 1993 ; Schmidt 1993b ; Schulz 1993 ; Staffeldt 1993; Wienert 1991), three tested combinations of hypericum and other herbal extracts (Ditzler 1992; Kniebel 1988; Steger 1985), and two compared hypericum extract with medications which are no longer considered adequate for depression (diazepam or bromazepam) (Kugler 1990a ; Warnecke 1986); one of these trials also was not explicitly randomized. Due to the new exclusion criterion, we excluded 17 trials not restricted to patients with major depression. Fifteen had been included in the previous version of the review (Halama 1991; Harrer 1991; Hoffmann 1979; Hübner 1993; König 1993; Lehrl 1993; Osterheider 1992; Quandt 1993; Reh 1992; Schlich 1987; Schmidt 1989; Schmidt 1993; Sommer 1994; Vorbach 1994; Winkel 2000) while two were not (Agrawal 1994 , for which it had not been possible to obtain a full copy and Gu 2001, which was newly identified in the update searches). Finally, we excluded one previously included trial as the standard antidepressant treatment was far below recommended dosages (30 mg Amitriptyline daily; Bergmann 1993).

Risk of bias in included studies

The majority of the trials were of high quality. The median quality scores were 5 (out of 5, range 2 to 5) for the Jadad scale and 4.5 (out of 6; range 2 to 6) for the IV scale (see quality rating of the single trials in the Characteristics of included studies).

Sequence generation/allocation concealment

The information on how the random sequence was generated was reported or provided on request for 18 trials (in all cases a computer program). Twenty two trials reported an adequate method of allocation concealment (most often consecutively numbered medication).

Blinding

All trials were described as double-blind, but only one trial reported that blinding was tested (HDTSG 2002). In this three-armed trial (hypericum vs. sertraline vs. placebo) about a third (as expected by chance alone) of guesses made by physicians were correct for hypericum and placebo patients, but in 66% of sertraline patients ($p = 0.001$).

Incomplete outcome data

In some trials attrition rates were high (for example, Fava 2005; see Characteristics of included studies). All placebo-controlled trials included an intent to treat analysis.

Effects of interventions

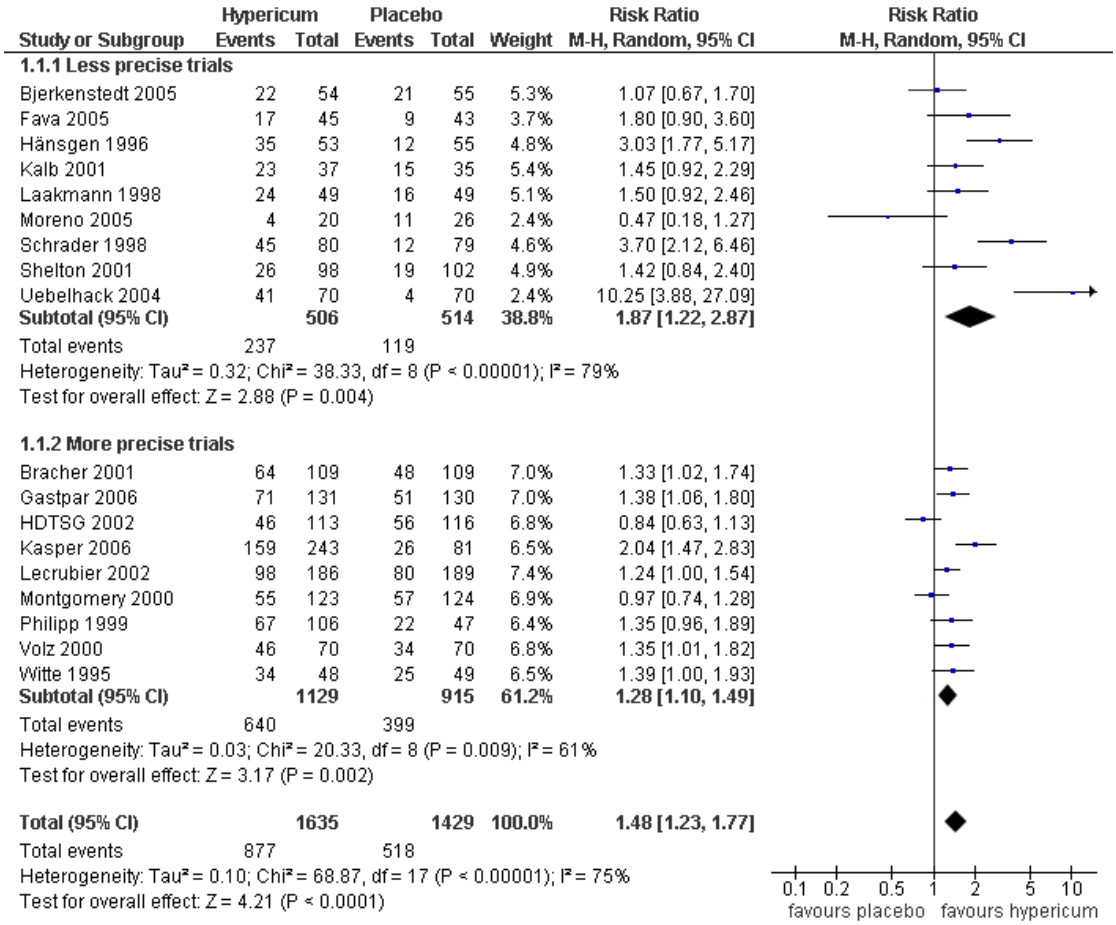
Comparison 1: Hypericum extracts versus placebo

1. Effectiveness

a) Responder analyses

Sixteen of the 18 placebo-controlled trials reported the number of patients classified as responders based on score reduction on the HAMD scale, one trial reported response according to the MADRS scale (Bracher 2001), and one trial only reported the proportion of patients rated at least as "improved" for the CGI (Volz 2002). Patients receiving hypericum extracts were significantly more likely to be responders (RR = 1.48; 95%CI 1.23 to 1.77; see comparison 1.1 and Figure 1) but study results were highly heterogeneous (I^2 = 75%). Effects in favour of hypericum extracts were less pronounced in more precise trials (RR = 1.28; 95%CI 1.10 to 1.49) compared to less precise trials (RR = 1.87; 95%CI 1.22 to 2.87) but heterogeneity was still strong in both subgroups (I^2 = 61% and 79%, respectively).

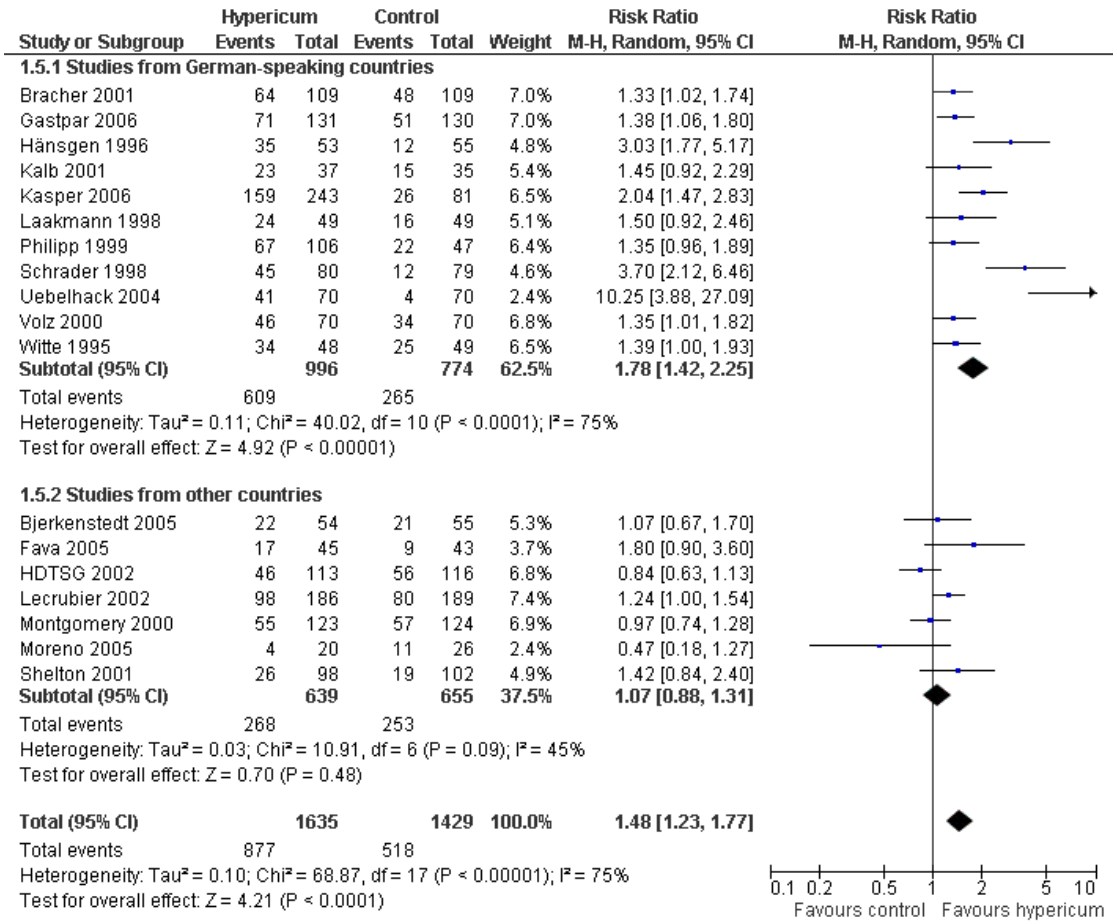
Figure 1. Forest plot of comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, outcome: 1.1 Responder - grouped by precision - primary analysis.



Findings were similar if response rates were based only on the trials reporting response according to the HAMD scale, or on the CGI (see comparisons 1.2 and 1.3). If trials investigating defined extracts were analysed separately (subgroups of trials testing the same extracts; see comparison 1.4), heterogeneity was strong in 3 of 4 subgroups.

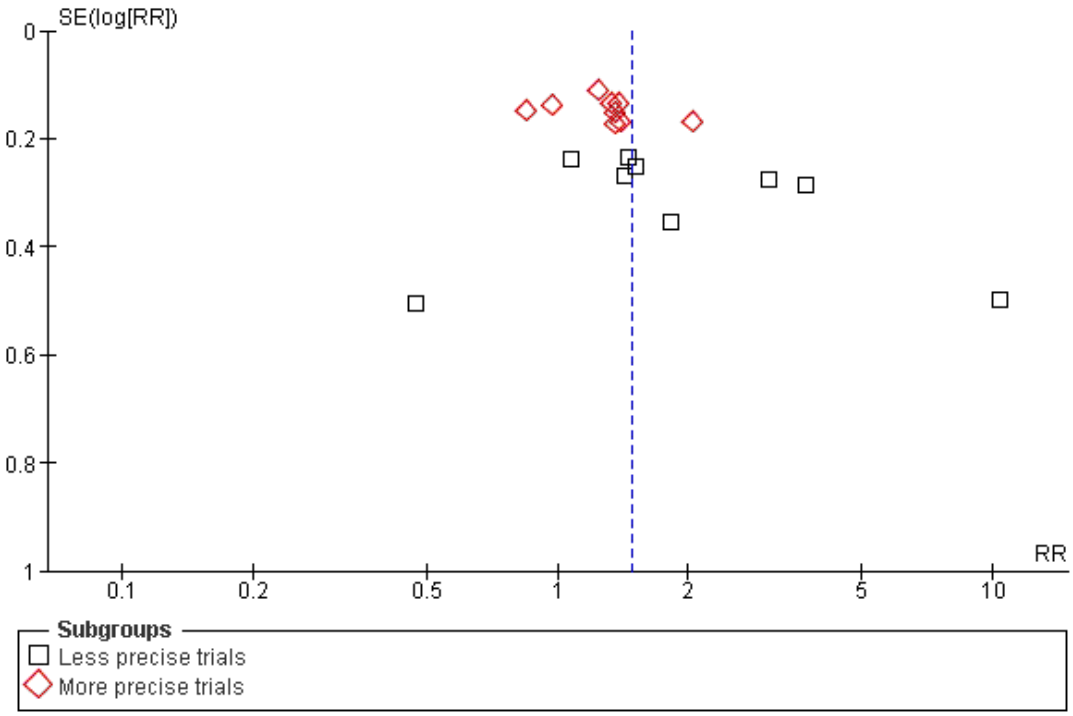
Trials from German-speaking countries reported more positive findings than trials from other countries (RR = 1.78; 95%CI 1.42 to 2.25 vs.1.07; 95% CI 0.88 to 1.31, respectively; see comparison 1.5 and Figure 2). Six trials reported remission rates. These were significantly higher in patients receiving hypericum extracts than in those receiving placebo (RR = 2.77; 95%CI 1.80 to 4.26; I^2 = 29%; see comparison 1.6).

Figure 2. Forest plot of comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, outcome: 1.5 Responder among studies from German-speaking countries and other studies.



There was significant funnel plot asymmetry for the main responder analysis (coefficient = 2.19, p = 0.03; Figure 3).

Figure 3. Funnel plot of comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, outcome: 1.1 Responder - grouped by precision - primary analysis.



In univariable meta-regression analyses, country of origin (studies from German-speaking countries showing larger effect sizes; p = 0.002), precision (more precise studies showing smaller effects; p = 0.032) and baseline values (higher values associated with smaller effect sizes; p = 0.048) were significantly associated with effects sizes. In multiple analyses the association remained significant for country of origin (p = 0.035) and precision (p = 0.017) but became non-significant for HAMD baseline values. Altogether over half of the variance (R^2 = 0.51) could be explained by these three variables. Findings of meta-regression analyses are summarised in Appendix 1.

b) Analyses of depression scales

Analyses based on mean HAMD values yielded similar findings. At the completion of treatment HAMD values were 3.04 (95%CI 1.78 to 4.29) score points lower in hypericum groups compared to placebo groups, but there was strong heterogeneity (I^2 = 86%; see comparison 2.1). Effects over placebo were significant after 2 (comparison 2.2), 4 (comparison 2.3), 6 to 8 weeks of treatment (comparison 2.4), and for changes from baseline to end of treatment (comparison 2.5). Significant effects over placebo were also reported for the MADRS (comparisons 2.6, and 2.7). Studies from German-language countries reported much larger effects over placebo (weighted mean difference = 4.29, 95%CI 2.97 to 5.61 score points; comparison 2.8) than studies from other countries (MD = 0.77 score points, 95%CI 0.20 to 1.74 score points). There was no significant funnel plot asymmetry (coefficient in the analysis of HAMD values at completion of treatment = -2.12, p = 0.35).

In multiple meta-regression analysis, country of origin was significantly associated with effects size (larger effects in trials from German-speaking countries; p < 0.001) but not precision and HAMD baseline values (R^2 = 0.63; see Appendix 1).

The four trials reporting results for the patient-rated von Zerssen Depression Scale (D-S) showed a significant effect of hypericum extracts over placebo (comparison 2.9). Post-hoc analyses using available data from 12 placebo-controlled trials for a variety of self-rating instruments also confirmed analyses based on physician-rated outcomes. The pooled standardised mean difference (SMD) was -0.47 (95% CI -0.64 to -0.30; I^2 = 74%; see comparison 2.10). Trials from German-speaking countries again reported more favourable findings than trials from other countries (SMDs of -0.57 and -0.17 respectively; see comparison 2.11).

2. Safety

Primary outcome

The number of patients dropping out for adverse effects was similar among patients receiving hypericum extracts and placebo (OR = 0.92, 95%CI 0.45 to 1.88, I^2=0%; see comparison 3.1).

Secondary outcomes

The total number of patients dropping out and the number of patients dropping out for any reason were similar among patients receiving hypericum extracts and placebo (comparisons 3.2 and 3.3).

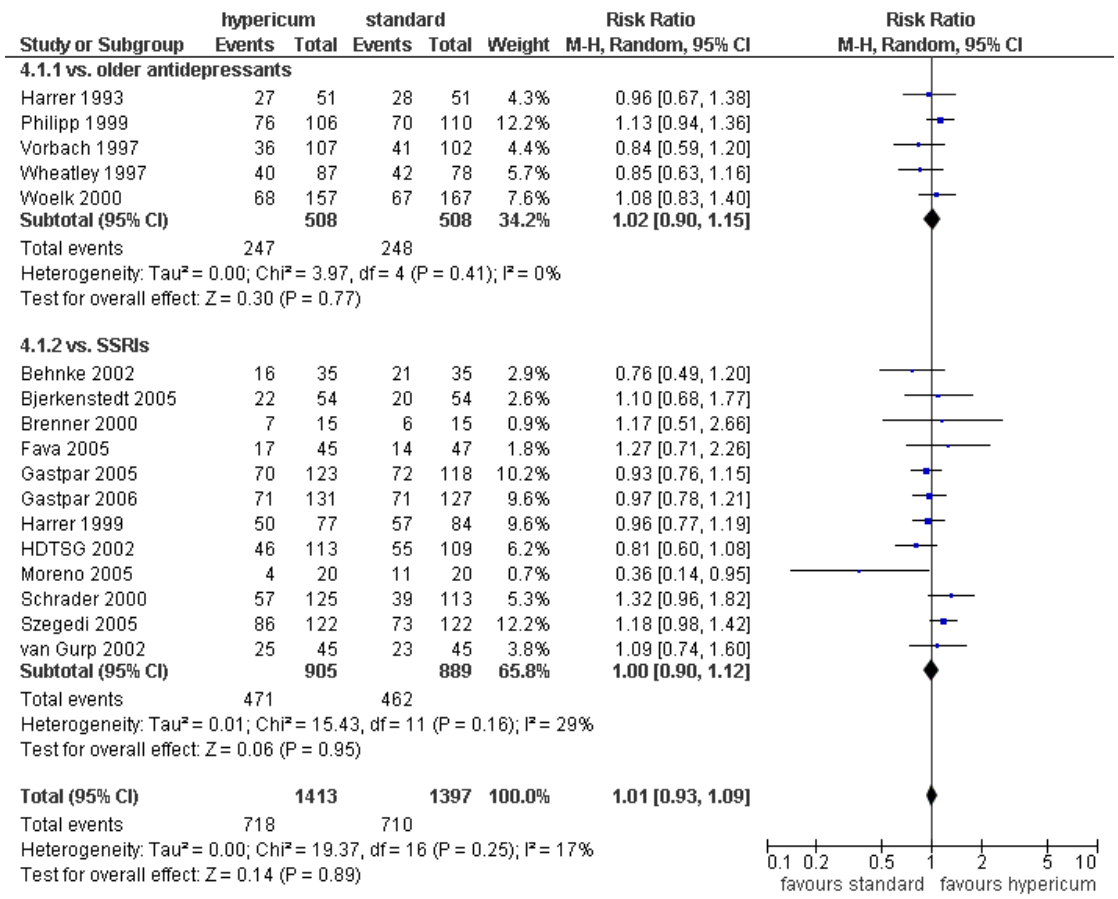
Comparison 2: Hypericum extracts versus standard antidepressants

1. Effectiveness

a) Responder analyses

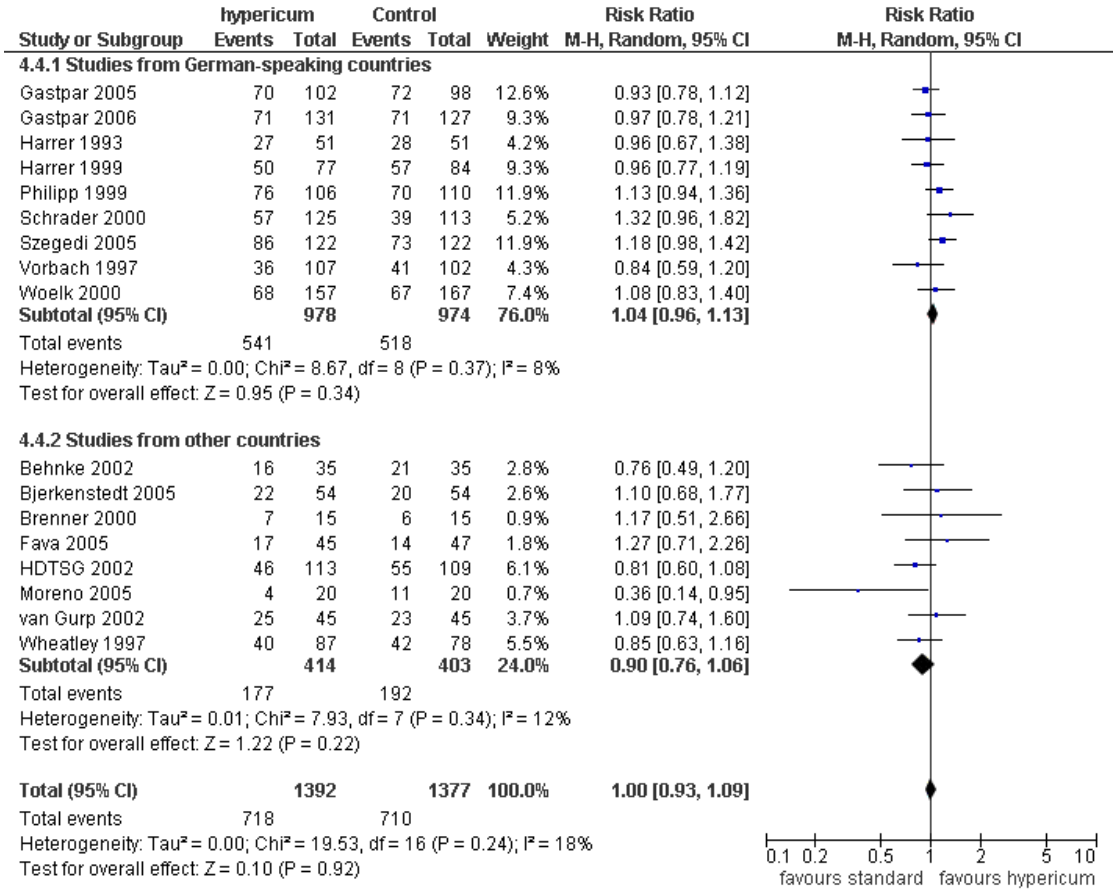
All 17 trials comparing hypericum extracts to standard antidepressant treatment reported the number of responders according to the HAMD score. Based on an intention to treat approach the pooled responder rate ratio was 1.01 for all 17 trials (95%CI 0.93 to 1.09; I^2 = 17%; see comparison 4.1 and Figure 4). For the five trials comparing hypericum extracts with older antidepressants, the pooled estimate was 1.02 (95%CI 0.90 to 1.15; I^2 = 0%), and 1.00 for the 12 trials with selective serotonine reuptake inhibitors (95%CI 0.90 to 1.12; I^2 = 29%).

Figure 4. Forest plot of comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, outcome: 4.1 Responder (intent to treat) - primary analysis.



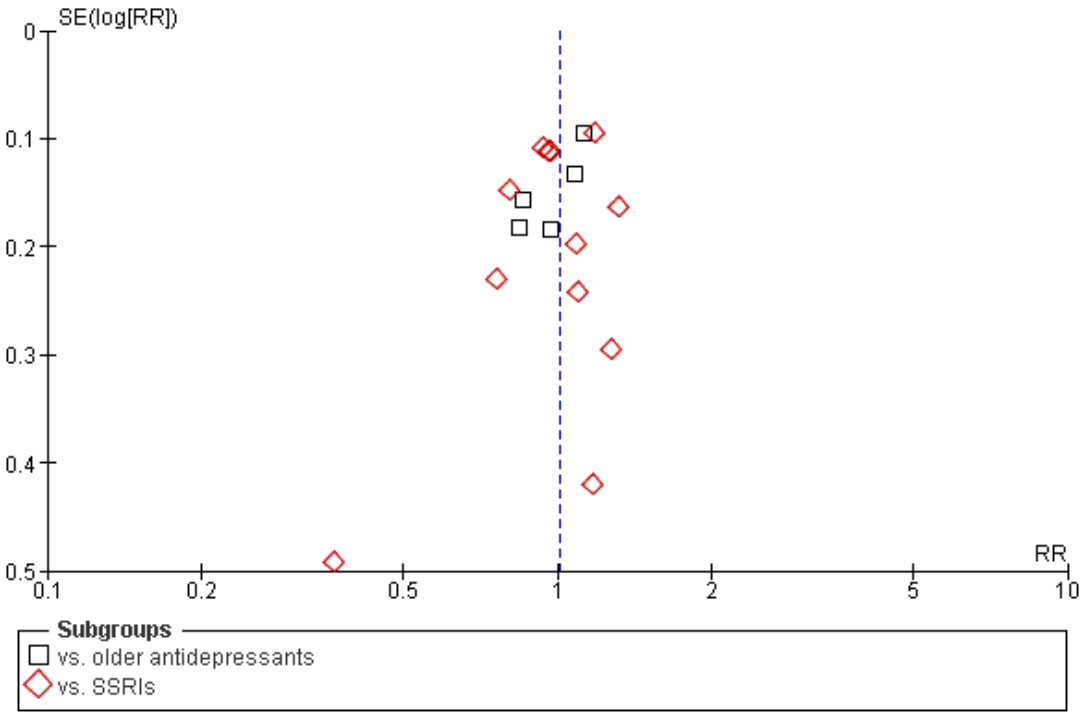
If analyses were based on per protocol data, the pooled responder rate ratio was 0.96 (95%CI 0.88 to 1.05; I^2 = 43%; see comparison 4.2). Analysis based on the CGI also found no relevant differences (RR = 1.01; 95%CI 0.94 to 1.09; I^2 = 24%; see comparison 4.3). In trials originating from German-speaking countries findings were slightly more favourable to hypericum than in trials from other countries (RR 1.04 and 0.90, respectively; see comparison 4.4 and Figure 5). In the four trials reporting remission rates the response rate ratio was 1.24 (95%CI 1.02 to 1.50; I^2 = 0%; see comparison 4.5).

Figure 5. Forest plot of comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, outcome: 4.4 Responder among studies from German-speaking studies and other studies.



The asymmetry coefficient for the main responder analysis was -1.07 (p = 0.09; see funnel plot in Figure 6). In univariable meta-regression analysis, there was a significant association between country of origin and response (trials from German-speaking countries favouring hypericum; p = 0.037). In the multivariable meta-regression analysis, none of the three tested predictors proved significant (R^2 = 0.24).

Figure 6. Funnel plot of comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, outcome: 4.1 Responder (intent to treat) - primary analysis.



b) Analyses of depression scales

Analyses based on HAMD scores confirmed the findings of the responder analysis (see comparisons 5.1 to 5.5, and 5.8). Analyses of MADRS- and D-S-values are difficult to interpret, as only few trials reported these outcomes (see comparisons 5.6, 5.7, 5.9). There was no funnel plot asymmetry (0.30, p = 0.73). In the multivariable meta-regression analysis, trials with higher HAMD baseline values showed less favourable results (p = 0.010), while country of origin and precision had no significant influence (R^2 = 0.44).

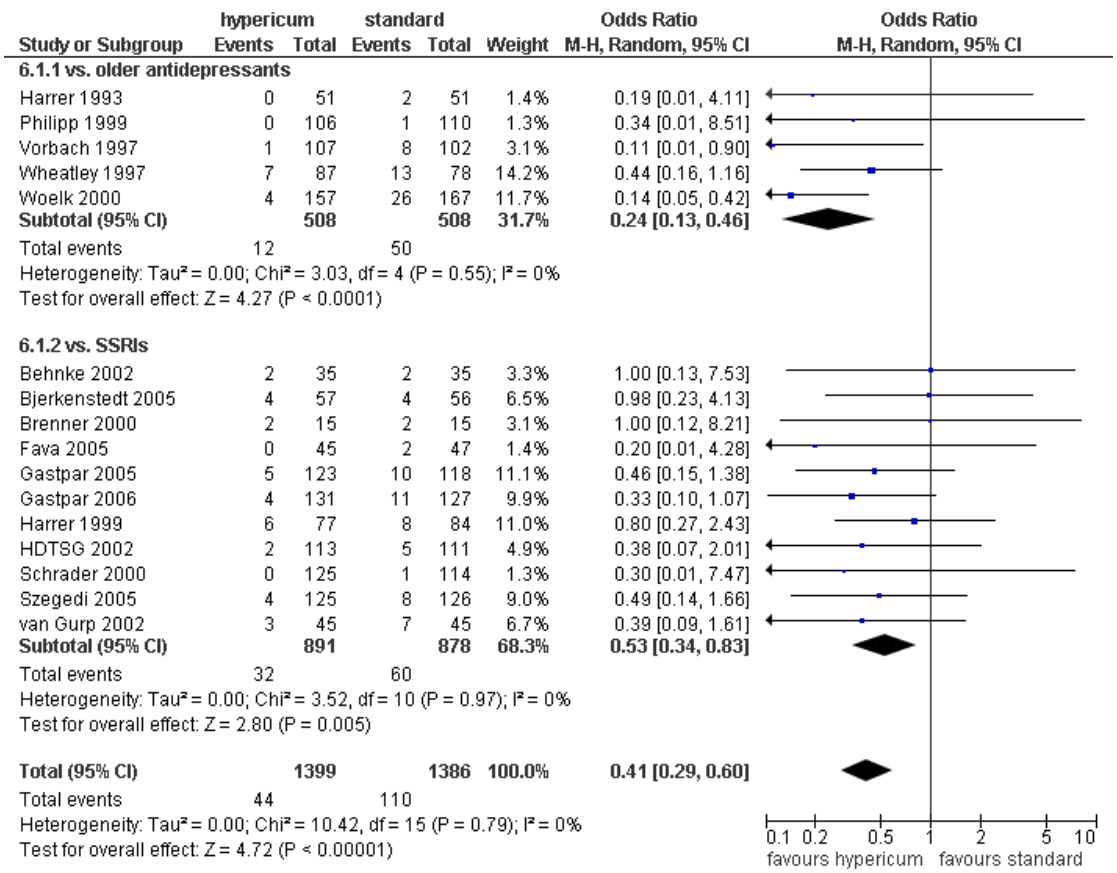
Again, post-hoc analyses using available data for a variety of self-rating instruments from 10 trials comparing hypericum extracts and standard antidepressants confirmed analyses based on physician-rated outcomes. The pooled SMD was 0.01 (95%CI -0.13 to 0.15; I^2 = 43%; see comparison 5.10). The pooled SMDs in trials from German-speaking countries was -0.02 compared to 0.10 in trials from other countries (comparison 5.11).

2. Safety

Primary outcome

Patients allocated to hypericum extracts were less likely to drop out from studies due to adverse effects than patients allocated to older standard antidepressants (OR = 0.24; 95%CI 0.13 to 0.46; I^2 = 0%) or to SSRIs (OR = 0.53; 95%CI 0.34 to 0.83; I^2 = 0%; see comparison 6.1 and Figure 7).

Figure 7. Forest plot of comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants, outcome: 6.1 Number of patients discontinuing treatment/dropping out due to adverse/side effects.



Secondary outcomes

Attrition for any reason was significantly lower for hypericum extracts compared to older antidepressants (OR = 0.67; 95%CI 0.47 to 0.95; I^2 = 0%), but not compared to SSRIs (OR = 0.83; 95% 0.63 to 1.08; I^2 = 0%; see comparison 6.2).

The number of patients reporting side effects was significantly higher in patients receiving older standard antidepressants (OR = 0.39; 95% 0.30 to 0.50; I^2 =0%); compared to SSRIs, the difference just missed significance (OR = 0.70; 95%CI 0.49 to 1.00; I^2 =57%; see comparison 6.3).

DISCUSSION

Summary of main results

Overall, the findings from newer trials seem to corrobate the evidence in favour of hypericum extracts. The available data suggest that the hypericum extracts tested in the included trials a) are superior to placebo in patients with major depression; b) are similarly effective as standard antidepressants; c) and have less side effects than standard antidepressants. There are two issues which complicate the interpretation of our findings: 1) While the influence of precision on study results in placebo-controlled trials is less pronounced in this updated version of our review compared to the previous version (Linde 2005a), results from more precise trials still show smaller effects over placebo than less precise trials. 2) Results from German-language countries are considerably more favourable for hypericum than trials from other countries.

Interpretation of the findings and limitations

For this update we excluded for the first time from our review all trials which were not restricted to patients with major depression. This does not mean that we believe that major depression is necessarily the only or best indication for hypericum extracts. Some authors argue that patients with signs of atypical depression might be particularly suited for treatment with hypericum extracts (Murck 2002 , Murck 2005), the National Institutes of Health are currently funding a trial in patients with minor depression (www.clinicaltrials.gov, identifier NCT00048815), and the findings from older trials (some of which, however, seem methodologically questionable) not restricted to patients with major discussion were very positive (Linde 1998). As pointed out in the introduction, we now focus on major depression to make our review better comparable to overviews on standard antidepressants, to have a more comparable set of studies for analysis, and also because almost all new trials of hypericum extracts are on this indication.

In spite of the tightened inclusion criteria, the findings of the placebo-controlled trials are still quite heterogeneous. In the trials a variety of hypericum extracts has been tested and daily doses cover a wide range. Differences in the interventions might contribute to some extent to the observed heterogeneity, but they do not seem to be a major factor. In three of four subgroup analyses of single extracts there was strong heterogeneity; for one extract the 95% confidence intervals of the two available trials (Gastpar 2006; Uebelhack 2004) did not even overlap, indicating that the results are hardly compatible. However, some of the factors leading to considerable heterogeneity between study findings could be identified. Considering country of origin, precision, and baseline depression severity of included patients explained 50 to 60 percent of the variance between trial results in comparisons with placebo and 20 to 40 percent in comparison with standard antidepressants. Nevertheless, it has to be stated that meta-regression analyses are (even if a priori defined) entirely of observative nature. Findings on the association of baseline depression severity and effect size estimates may be biased through structural dependence and regression to the mean, and thus should be interpreted with caution (Higgins 2008). Furthermore, inferences drawn from a meta-regression analysis on aggregate data may differentiate from inferences drawn from a meta-regression analysis on individual data (Deeks 2006; e.g. 'ecological bias').

The finding that more precise placebo-controlled trials yielded less positive results than less precise trials could indicate publication bias (trials with positive results are more likely to be published than trials with negative results) or bias within studies (smaller trials with less rigorous methods yielding overoptimistic results). We cannot rule out, but doubt, that selective publication of overoptimistic results in small trials strongly influences our findings. There is some evidence that "negative" trials without demonstrable differences between extracts and placebo were published less often as full articles than trials with "positive" findings. Our extensive searches identified three "negative" trials that were only published as abstracts or theses. Two that were conducted in the early 1990s (these were included in earlier versions of this review Linde 1996; Linde 1998; Linde 2005a) involved patients without documented major depression (König 1993; Osterheider 1992), and one that was conducted in the late 1990s involved patients with major depression (Montgomery 2000). One positive trial included in our last update, but now excluded, was published as an abstract and as a chapter in a not widely available book (Winkel 2000). One comparably large, positive trial (Bracher 2001) has been published only in a short report as a supplement to a German medical newspaper. This trial is an example that sponsors or manufacturers of herbal medicines sometimes have very limited interest in a major publication if their trial includes a new aspect (in that case a once daily dosage), as there is no patent protection for herbal extracts and results can be exploited by competitors, too. We suspect that there are few additional relevant unpublished trials. Few manufacturers of hypericum extracts sponsor research trials, and the five manufacturers whose products were tested in most of the trials told us they had (with the exception of one smaller negative trial) no other unpublished trials that possibly met our inclusion criteria. Through personal communication we were informed that there are at least one or two unpublished negative trials on tea

preparations of Hypericum. However, tea preparations are phytochemically very different from alcoholic extracts and have to be evaluated separately.

We found that the quality of the majority of trials was adequate, and we detected no systematic differences in design aspects known to be potential sources of bias. All trials were double-blind. Though adequacy of blinding was not formally assessed in most trials, achieving similarity between hypericum and placebo preparations is not particularly difficult. All trials were randomised, and most concealed allocation assignments by using consecutively numbered identical medication containers. Reported drop-out rates were low in the majority of trials. Investigators involved in older trials may have had less training and/or experience with diagnostic standards and rating scales for depressive symptoms (Shelton 2001), but this issue, if true, is likely to affect generalisation of findings rather than internal validity. Finally, though we found no systematic differences in major factors generally related to trial quality, our subjective judgement was that larger trials tended to be of better quality than smaller trials. The dosages of standard antidepressants were (with two exceptions) within the range recommended in current guidelines (e.g., Härter 2003), but at the lower limit.

Our finding that studies from German-speaking countries yielded more favourable results than trials performed elsewhere is difficult to interpret. As our analyses are partly data-driven, they must be considered cautiously. However, the consistency and extent of the observed association suggest that there are important differences in trials performed in different countries. One possibility is that studies performed in German-speaking countries with a long history of hypericum prescription by physicians enrolled slightly different patients in spite of similar inclusion criteria. With one exception (the extremely positive trial by Uebelhack 2004 performed in a research clinic of a contract research organisation), all German studies recruited patients in private practices, while a number of trials from other countries were performed in academic research settings or hospital outpatient units. Depression with atypical or reversed vegetative features might be present more often in primary care outpatient populations (Murck 2005).The trend that trials with higher HAMD baseline values reported slightly less favourable results also suggests that effectiveness of hypericum extracts might differ between subgroups of depressive patients. While we did not systematically investigate this issue, it seems to us that the trials from countries other than Germany might be more often investigator-initiated. A closer link of trial planning, performance and analysis with manufacturer interests might influence study findings. This could result possibly in true bias, but also in conditions making a true positive outcome more likely. For example, for at least three trials (Kasper 2006; Schrader 1998; Uebelhack 2004) with large effects performed in German-speaking countries authors or sponsors reported in the publication or in personal communications that contact times and interaction with patients were limited to minimise placebo response rates. Increasing placebo group response rates due to the intensive care and monitoring in antidepressant trials are considered by some researchers as a potential reason for the problem to show specific effects (Posternak 2007). One could also speculate whether unblinding might lead German physicians (who often use hypericum extracts in their usual practice) to give more positive ratings and (the possibly more sceptic) colleagues from elsewhere to more negative ratings. However, as hypericum extracts have no characteristic side effects such a problem seems only relevant in comparisons with standard antidepressants.

Potential biases in the review process

The work for the first version of this review started in 1993 and three previous versions are available (Linde 1996; Linde 1998; Linde 2005a). During that period a large number of new trials became available, diagnostic classifications used for including patients into studies have changed and the quality of trials has improved. In parallel the methods of our review were adapted. The changes over time make it difficult to report our searches and their results in a consistent and transparent manner. The way how we approached authors/sponsors for obtaining missing information and the contents of inquiries were not fully systematic and have changed over time. This could imply that additional data necessary for some secondary analyses were obtained for a selected subset of studies. However, data for the main analysis were available for all or almost all trials, therefore, major biases seem highly unlikely. A potential source of bias in the responder analyses could be slightly variable responder definitions in the primary studies. Response according to the HAMD was either defined as at least 50% reduction, a HAMD score < 10 (or 11) after treatment, at least one or the combination of both. Whether these definitions were truly made a priori in each study could not be assessed. Decisions on the inclusion of subgroup analyses (for example, regarding precision or country effects) for updates were driven by findings in previous versions of the review. Therefore, these analyses must be interpreted with caution. Publication and small study bias have been discussed in the previous section.

AUTHORS' CONCLUSIONS

Implications for practice

In older systematic reviews and meta-analyses of hypericum extracts (Ernst 1995; Linde 1996; Linde 1998; Kim 1999; Gaster 2000; Williams 2000) the findings of the included studies were mostly positive, but reviewers drew cautious conclusions due to methodological limitations. The quality of trials on average clearly improved over recent years. However, study findings became more often contradictory, and in the last version of our review larger trials restricted to patients with major depression showed only minimal effects over placebo (Linde 2005a). With the addition of

several new, partly large trials, the cumulative evidence now suggests that hypericum extract have a modest effect over placebo in a similar range as standard antidepressants (Kirsch 2008; Turner 2008). The direct comparisons with older antidepressants and selective serotonine reuptake inhibitors seem to confirm this impression. The available clinical trials now also show that hypericum extracts have fewer side effects than both older antidepressants and selective serotonine reuptake inhibitors. This would imply that an attempt of treating mild to moderate major depression with one of the hypericum preparations positively tested in clinical trials is clearly justified. However, the differences in the findings from different countries make clear-cut recommendations difficult. The evidence for severe major depression is still insufficient to draw conclusions.

Many patients buy St John's wort products from health food stores and might not disclose this to their physicians (Smith 2004). Such uncontrolled use is problematic as serious interactions can occur with a number of frequently used drugs (Ernst 1999; Hammerness 2003 ; Knüppel 2004 ; Whitten 2006). Therefore, physicians should regularly ask their patients about hypericum intake. However, it must be kept in mind that drug interactions are not a problem unique to hypericum extracts, but also common for standard antidepressants (Nieuwstraten 2006).

It has to be emphasised that the quality of hypericum preparations can differ considerably. The composition of a product depends on the raw plant material used, the extraction process, and the solvents. In consequence, the amounts of bioactive constituents in different products can vary enormously. A recent study has shown that a number of products available on the German market contain only minor amounts of bioactive constitutents (Wurglics 2003). The hypericum extracts tested in clinical trials have to be considered high quality products. Results obtained with these extracts cannot be extrapolated directly to other products. In our meta-analysis, the type of extract did not contribute to the explanation of heterogeneity. This does not mean, however, that all products tested in the available trials are equally effective. Standardisation of a product on a defined component (for example, hyperforin or hypericin) does not resolve the problem, as currently the exact mechanism for the antidepressant effects of hypericum extracts is still unclear, and available research indicates that several components are relevant. The findings of this review most likely apply to products (using ethanol 50 to 60% or methanol 80% for extraction from dried plant material) with daily extract dosages of 500 to 1200 mg with a ratio of raw material to extract of 3-7:1.

Implications for research

There is a clear need to investigate the reasons for the differences in findings from trials originating from German-speaking countries and those from other countries. Mulitnational trials would seem desirable, but it is unlikely that there will be funding for such studies in the near future. Individual patient data meta-analysis of existing trials could be a possible tool to investigate predictors of treatment response in a more accurate manner. The authors will try to obtain such data from researchers and/or sponsors.

ACKNOWLEDGEMENTS

The authors would like to thank

- the researchers and sponsors who provided additional information on trials

- Eva Israel and Karsten Schmidt who helped with assessing eligibility and doing extractions

- Cindy Mulrow and Matthias Egger who were involved in previous versions of the review

REFERENCES

References to studies included in this review

Behnke 2002 *{published data only}*
Behnke K, Jensen GS, Graubaum HJ, Gruenwald J. [Hypericum perforatum versus fluoxetine in the treatment of mild to moderate depression]. *Advances in Therapy* 2002;**19**:43–52.

Bjerkenstedt 2005 *{published data only}*
Bjerkenstedt L, Edman GV, Alken RG, Mannel M. Hypericum extract LI 160 and fluoxetine in mild to moderate depression. A randomized, placebo-controlled multi-center study in outpatients. *European Archives of Psychiatry and Clinical Neuroscience* 2005;**255**:40–7.

Bracher 2001 *{published and unpublished data}*
Bracher A. Trial on HYP611 confirms effectiveness [HYP611–Studie bestätigt Wirksamkeit]. *Ärztliche Praxis* 2001;**51 (suppl.)**:3–4.

Brenner 2000 *{published data only}*
Brenner R, Azbel V, Madhusoodanan S, Pawlowska M. Comparison of an extract of hypericum (LI 160) and sertraline in the treatment of depression: a double-blind, randomized pilot study. *Clinical Therapeutics* 2000;**22**:411–9.

Brenner R, Madhusoodanan S, Pawlowska M. Efficacy of continuation treatment with hypericum perforatum in depression. *Journal of Clinical Psychiatry* 2002;**63**:455.

Fava 2005 *{published data only}*
Fava M, Alpert J, Nierenberg AA, Mischoulon D, Otto MW, Zajecka

J, Murck H, Rosenbaum JF. A double-blind, randomized trial of St John's wort, fluoxetine, and placebo in major depressive disorder. *Journal of Clinical Psychopharmacology* 2005;**25**:441–7.

Gastpar 2005 *{published and unpublished data}*
Gastpar M, Singer A, Zeller K. Efficacy and tolerability of hypericum extract STW3 in long-term treatment with a once-daily dosage in comparison with sertraline. *Pharmacopsychiatry* 2005;**38**:78–87.

Gastpar 2006 *{published and unpublished data}*
Gastpar M, Singer A, Zeller K. Comparative efficacy and safety of a once-daily dosage of hypericum extract STW3-VI and citalopram in patients with moderate depression: a double-blind, randomised, multicentre, placebo-controlled study. *Pharmacopsychiatry* 2006;**39**: 66–75.

Harrer 1993 *{published and unpublished data}*
Harrer G, Hübner WD, Podzuweit H. Effectiveness and tolerance of the hypericum extract LI 160 compared to maprotiline: a multicenter double-blind study. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S24–8.

Harrer G, Hübner WD, Podzuweit H. Wirksamkeit und Verträglichkeit des Hypericum-Präparates LI 160 im Vergleich mit Maprotilin. *Nervenheilkunde* 1993;**12**:297–301.

Harrer 1999 *{published data only}*
Harrer G, Schmidt U, Kuhn U, Biller A. Comparison of equivalence between the St. John's wort extract LoHyp-57 and fluoxetine. *Arzneimittel-Forschung/Drug Research* 1999;**49**:289–96.

HDTSG 2002 *{published and unpublished data}*
* Hypericum Depression Trial Study Group. [Effect of Hypericum perforatum (St John's wort) in major depressive disorder. A randomized controlled trial]. *JAMA* 2002;**287**:1807–14.

Vitiello B, Shader RI, Parker CB, Ritz L, Harlan W, Greenblatt DJ, Gadde KM, Krishnan KRR, Davidson JRT. Hyperforim plasma level as a marker of treatment adherence in the National Institutes of Health Hypericum Depression Trial. *Journal of Clin Psychopharmacology* 2005;**25**:243–9.

Hänsgen 1996 *{published and unpublished data}*
* Hängsen KD, Vesper J. Antidepressive Wirksamkeit eines hochdosierten Hypericum-Extraktes. *Muenchener medizinische Wochenschrift* 1996;**138**(3):29–33.

Hänsgen KD, Vesper J, Ploch M. Multizentrische Doppelblindstudie zur antidepressiven Wirksamkeit des Hypericumextraktes LI 160. *Nervenheilkunde* 1993;**12**:285–9.

Vesper J, Ploch M. Multi-center double-blind study examining the antidepressant effectiveness of the hypericum extract LI 160. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S15–8.

Kalb 2001 *{published data only}*
Kalb R, Trautmann-Sponsel RD, Kieser M. Efficacy and tolerability of hypericum extract WS 5572 versus placebo in mildly depressed patients. *Pharmacopsychiatry* 2001;**34**:96–103.

Kasper 2006 *{published and unpublished data}*
Kasper S, Anghelescu IG, Szegedi A, Dienel A, Kieser M. Superior efficacy of St John's wort extract WS 5570 compared to placebo in patients with major depression: a randomized, double-blind, placebo-controlled, multi-center trial [ISRCTN77277298]. *BMC Medicine* 2006;**4**:14.

Laakmann 1998 *{published data only}*
Laakmann G, Dienel A, Kieser M. Clinical significance of hyperforin for the efficacy of hypericum extracts on depressive disorders of different severities. *Phytomedicine* 1998;**5**:435–42.

Laakmann G, Schüle C, Baghai T, Kieser M. St John's wort in mild and moderate depression: the relevance of hyperforin for the clinical efficacy. *Pharmacopsychiatry* 1998;**31 (suppl)**:54–9.

Lecrubier 2002 *{published and unpublished data}*
Lecrubier Y, Bech P. The Ham D6 is more homogeneous and as sensitive as the Ham D17. *European Psychiatry* 2007;**22**:252–5.

* Lecrubier Y, Clerc G, Didi R, Kieser M. Efficacy of St. John's wort extract WS 5570 in major depression: a double-blind, placebo-controlled trial. *American Journal of Psychiatry* 2002;**159**:1361–6.

Montgomery 2000 *{published and unpublished data}*
Montgomery SA, Hübner WD, Grigoleit HG. Efficacy and tolerability of St. John's wort extract compared with placebo in patients with a mild to moderate depressive disorder. *Phytomedicine* 2000;**7 (suppl II)**:107.

Moreno 2005 *{published data only}*
Moreno RA, Teng CT, de Almeida KM, Tavares H jr. Hypericum perforatum versus fluoxetine in the treatment of mild to moderate depression: a randomized double-blind trial in a Brazilian sample. *Revista Brasileira de Psiquiatria* 2005;**28**:29–32.

Philipp 1999 *{published and unpublished data}*
Philipp M, Kohnen R, Hiller KO. Hypericum extract versus imipramine or placebo in patients with moderate depression: randomized multicentre study of treatment for 8 weeks. *BMJ* 1999;**319**: 1534–9.

Schrader 1998 *{published and unpublished data}*
Schrader E, Meier B, Brattström A. Hypericum treatment of mild-moderate depression in a placebo-controlled study. A prospective, double-blind, randomized, placebo-controlled, mulitcentre study. *Human Psychopharmacology* 1998;**13**:163–9.

Schrader 2000 *{published and unpublished data}*
Schrader E. Equivalence of St John's wort (Ze 117) and fluoxetine: a randomized, controlled study in mild-moderate depression. *International Clinical Psychopharmacology* 2001;**15**:61–8.

Shelton 2001 *{published and unpublished data}*
Gelenberg AJ, Shelton RC, Crits-Christoph P, Keller MB, Dunner DL, Hirschfeld RMA, Thase MAE, Russel JM, Lydiard RB, Gallop RJ, Todd L, Hellersteim DJ, Goodnick PJ, Keitner GI, Stahl SM, Halbreich U, Hopkins KS. The effectiveness of St. John's wort in major depressive disorder: a naturalistic phase 2 follow-up in which non-responders were provided alternate medication. *Journal of Clinical Psychiatry* 2004;**65**:1114–9.

* Shelton RC, Keller MB, Gelenberg A, Dunner DL, Hirschfeld R, Thase ME, et al.Effectiveness of St. John's wort in major depression. A randomized controlled trial. *JAMA* 2001;**285**:1978–86.

Szegedi 2005 *{published and unpublished data}*
Anghelescu IG, Kohnen R, Szegedi A, Klement S, Kieser M. Comparison of hypericum extract WS 5570 and paroxetine in ongoing treatment after recovery from an episode of moderate to severe depression: results from a randomized multicenter study. *Pharmacopsychiatry* 2006;**39**:213–9.

* Szegedi A, Kohnen R, Dienel A, Kieser M. Acute treatment of moderate to severe depression with hypericum extract WS 5570 (St John's wort): randomised controlled double blind non-inferiority trial versus paroxetine. *BMJ* 2005;**330**:503–6.

Uebelhack 2004 *{published and unpublished data}*
Uebelhack R, Gruenwald J, Graubaum HJ, Busch R. Efficacy and tolerability of hypericum extract STW 3-VI in patients with moderate depression: a double-blind, randomized, placebo-controlled clinical trial. *Advances in Therapy* 2004;**21**:265–75.

van Gurp 2002 *{published and unpublished data}*
van Gurp G, Meterissian GB, Haiek LN, McCusker J, Bellavance F. St John's wort or sertraline? Randomized controlled trial in primary care. *Canadian Family Physician* 2002;**48**:905–12.

Volz 2000 *{published data only}*
Volz HP, Eberhardt R, Grilll G. Efficacy and tolerance of the St John's wort extract D-0496 in mild to moderate depression - a placebo-controlled, double-blind 6-week trial [Wirksamkeit und Verträglichkeit des Johanniskrautextraktes D–0496 bei leichten bis mittelschweren depressiven Episoden. Plazebokontrollierte Doppelblindstudie über 6 Wochen]. *Nervenheilkunde* 2000;**19**:401–5.

Vorbach 1997 *{published and unpublished data}*
Vorbach EU, Arnoldt KH, Hübner WD. Efficacy and tolerability of St. John's wort extract LI 160 versus imipramine in patients with severe depressive episodes according to ICD-10. *Pharmacopsychiatry* 1997;**30**(suppl.2):81–5.

Wheatley 1997 *{published and unpublished data}*
Wheatley D. LI 160, an extract of St. John's wort, versus amitriptyline in mildly to moderately depressed outpatients - a controlled 6-week clinical trial. *Pharmacopsychiatry* 1997;**30**(suppl.2):77–80.

Witte 1995 *{published and unpublished data}*
Witte B, Harrer G, Kaptan T, Podzuweit H, Schmidt U. Behandlung depressiver Verstimmungen mit einem hochkonzentrierten Hypericumpräparat - eine multizentrische plazebokontrollierte Doppelblindstudie. *Fortschritte der Medizin* 1995;**113**:404–8.

Woelk 2000 *{published data only}*
Woelk H for the Remotiv/Imipramine Study Group. Comparison of St John's wort and imipramine for treating depression: randomised controlled trial. *BMJ* 2000;**321**:536–9.

References to studies excluded from this review

Agrawal 1994 *{published data only}*
Agrawal A, Dixit SP, Dubey GP, Gupta BS. Clinical evaluation of anti-depressant properties of basant (hypericum perforatum). *Pharmacopsychoecologia* 1994;**7**:253–6.

Albertini 1986 *{published data only}*
Albertini H. Evaluation d'un traitement homeopathique de la nevralgie dentaire. In: Boiron J, Belon P, Hariveau E editor(s). *Recherche en homeopathie*. Lyon: Fondation francaise pour la recheche en homeopathie, 1986:75–7.

Bendre 1980 *{published data only}*
Bendre VV, Dharmadhikari SD. Arnica montana and hypericum in dental practice. *Hahnemann Gleanings* 1980;**47**:70–2.

Bergmann 1993 *{published and unpublished data}*
Bergmann R, Nübner J, Demling J. Behandlung leichter bis mittelschwerer Depressionen. *Therapiewoche Neurol Psychiatr* 1993;**7**: 235–40.

Bernhadt 1993 *{published data only}*
Bernhardt M, Liske E, Ebeling L. Hypericum perforatum in der Therapie leichter bis mittelschwerer Depressionen: Vergleich der antidepressiven Wirksamkeit von zwei unterschiedlichen Dosierungsschemata. Bonn, 5. Phytotherapie-Kongreß 1993.

Brockmöller 1997 *{published data only}*
Brockmöller J, Reum T, Bauer S, Kerb R, Hübner WD, Roots I. Hypericin and pseudohypericin: pharmacokinetics and effects on photosensitivity in humans. *Pharmacopsychiatry* 1997;**30**(suppl.2):94–101.

Czekalla 1997 *{published data only}*
Czekalla J, Gastpar M, Hübner WD, Jäger D. The effect of hypericum extract on cardiac conduction as seen in the electrocardiogram compared to that of imipramine. *Pharmacopschychiatry* 1997; **30**(supp.2):86–8.

Dittmer 1992 *{published data only}*
Dittmer TLJ. Die Behandlung von psychovegetativen Störungen mit Johanniskraut-Öl (Jukunda-Rot-Öl-Kapseln). *Naturheilpraxis mit Naturmedizin* 1992;**45**:118–22.

Ditzler 1992 *{published and unpublished data}*
Ditzler K, Gessner B, Schatton WFH, Willems M. Clinical trial on Neuropas versus placebo in patients with mild to moderate depressive symptoms: a placebo-controlled, randomised double-blind study. *Complementary Therapies in Medicine* 1994;**2**:5–13.

Ditzler K, Schatton W. Johanniskraut bei leichten bis mittelschweren Depressionen - Ergebnisse einer plazebo-kontrollierten Doppelblindstudie. *Heilkunst* 1992;**104**:263–70.

Gu 2001 *{published data only}*
Gu N, Li H, Shu L. Zhang H, Weng Z, Ma C, Wang M. Multicenter study of St John's wort extract in treatment of mild to moderate depression. *Chinese Journal of Clinical Pharmacy* 2001;**10**:271–4.

Halama 1991 *{published and unpublished data}*
Halama P. Wirksamkeit des Hypericum-Extraktes LI 160 bei 50 Patienten einer psychiatrischen Fachpraxis. *Nervenheilkunde* 1991;**10**: 305–7.

Halama P. Wirksamkeit des Johanniskrautextraktes LI 160 bei depressiver Verstimmung. *Nervenheilkunde* 1991;**10**:250–3.

Harrer 1991 *{published and unpublished data}*
Harrer G, Schmidt U, Kuhn U. "Alternative" Depressionsbehandlung mit einem Hypericum-Extrakt. *Therapiewoche Neurologie* 1991; **5**:710–6.

Herberg 1992 *{published data only}*
Herberg KW. Psychotrope Phytopharmaka im Test - Alternative zu synthetischen Psychopharmaka. *Therapiewoche* 1992;**44**:704–13.

Hoffmann 1979 *{published data only}*
Hoffmann J, Kühl ED. Therapie von depressiven Zuständen mit Hypericin. *Zeitschrift für Allgemeinmedizin* 1979;**55**:776–82.

Hottenrott 1997 *{published data only}*
Hottenrott K, Sommer HM, Lehrl S, Hauer H. Der Einfluss von Vitamin E und Johanniskraut-Trockenextrakt auf die Ausdauer-

leistungsfähigkeit von Wettkampfsportlern. Eine placebo-kontrollierte Doppelblindstudie mit Langstreckenläufern und Triathleten. *Deutsche Zeitschrift für Sportmedizin* 1997;**48**:22–7.

Häring 1996 *{published and unpublished data}*
Häring B, Hauns B, Hermann C, Hübner WD, Maier-Lenz H, Marschner N. A double-blind, placebo-controlled pilot study of LI 160 in combination with chemotherapy in patients with solid tumors. Abstracts of the 2nd International Congress on Phytomedicine. Munich, 1996:SL–88.

Hübner 1993 *{published and unpublished data}*
Hübner WD, Lande S, Podzuweit H. Behandlung larvierter Depressionen mit Johanniskraut. *Nervenheilkunde* 1993;**12**:278–280.

Hübner WD, Lande S, Podzuweit H. Hypericum treatment of mild depression with somatic symptoms. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S12–5.

Johnson 1992 *{published data only}*
Johnson D. Neurophysiologische Wirkungen von Hypericum im Doppelblindversuch mit Probanden. *Nervenheilkunde* 1991;**10**:316–7.

* Johnson D, Siebenhüner G, Hofer E, Sauerwein-Giese E, Frauendorf A. Einfluß von Johanniskraut auf die ZNS-Aktivität. *Neurologie und Psychiatrie* 1992;**6**:436–44.

Johnson 1993 *{published data only}*
Johnson D, Ksciuk H, Woelk H, Sauerwein-Giese E, Frauendorf A. Effetcs of hypericum extract LI 160 compared with maprotiline on resting EEG and evoked potentials in 24 volunteers. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S44–6.

Johnson D, Ksciuk H, Woelk H, Sauerwein-Giese E, Frauendorf A. Wirkungen mit Johanniskraut-Extrakt LI 160 im Vergleich mit Maprotilin auf Ruhe-EEG und evozierte Potentiale bei 24 Probanden. *Nervenheilkunde* 1993;**12**:328–30.

Kniebel 1988 *{published and unpublished data}*
Hiller KO, Rahlfs V. Therapeutische Äquivalenz eines hochdosierten Phytopharmakons mit Amitriptylin bei ängstlich-depressiven Verstimmungen - Reanalyse einer randomisierten Studie unter besonderer Beachtung biometrischer und klinischer Aspekte. *Forschende Komplementärmedizin* 1995;**2**:123–32.

* Kniebel R, Burchard JM. Zur Therapie depressiver Verstimmungen in der Praxis. *Zeitschrift für Allgemeinmedizin* 1988;**64**:689–96.

Kugler 1990a *{published data only}*
Kugler J, Weidenhammer W, Schmidt A, Groll S. [Therapie depressiver Zustände. Hypericum–Extrakt Steigerwald als Alternative zur Benzodiazepin–Behandlung]. *Zeitschrift für Allgemeinmedizin* 1990; **66 (suppl 1)**:21–9.

Kugler 1990b *{published data only}*
Kugler J, Schmidt A, Groll S, Weidenhammer W. Zur Pharmakodynamik eines Hypericum-Extraktes. Untersuchungen bei Patienten mit depressiven Zuständen im Vergleich zu Bromazepam und Placebo. *Zeitschrift für Allgemeinmedizin* 1990;**66**(suppl.):13–20.

König 1993 *{published data only}*
König CD. *Hypericum perforatum L. (gemeines Johanniskraut) als Therapeutikum bei depressiven Verstimmungszuständen - eine Alternative zu synthetischen Arzneimitteln*. University of Basel: Thesis, 1993.

Lehrl 1993 *{published and unpublished data}*
* Lehrl S, Woelk H. Ergebnisse von Messungen der kognitiven Leistungsfähigkeit bei Patienten unter der Therapie mit Johanniskraut. *Nervenheilkunde* 1993;**12**:281–4.

Lehrl S, Woelk H. Psychometrische Messung der Leistungkapazität unter antidepressiver Therapie mit Johanniskraut. *Nervenheilkunde* 1991;**10**:313–5.

Lenoir 1999 *{published data only}*
Lenoir S, Degenring FH, Saller R. [A double–blind randomised trial to investigate three different concentrations of a standardised fresh plant extract obtained from the shoot tips of hypericum perforatum L]. *Phytomedicine* 1999;**6**:141–6.

Li 2005 *{published data only}*
Li YM, Li GY, Wang XL, & Yu LF. Rehabilitation effect of Saint John's Wort extract on depression and myocardial function after coronary artery bypass grafting: A randomized grouping, placebo-control and blind evaluation. *Chinese Journal of Clinical Rehabilitation* 2005;**9** (28):38–9.

Maisenbacher 1995 *{published data only}*
Maisenbacher J, Schmidt U, Schenk N. Therapie mit Hypericum bei Angstzuständen. *Therapiewoche Neurologie Psychiatrie* 1995;**9**: 65–60.

Martinez 1993 *{published data only}*
Kasper S. Treatment of seasonal affective disorder (SAD) with hypericum extract. *Pharmacopsychiatry* 1997;**30**(suppl.2):89–93.

Martinez B, Kasper S, Ruhrmann B, Möller HJ. Hypericum in der Behandlung von saisonal abhängigen Depressionen. *Nervenheilkunde* 1993;**12**:302–7.

Martinez B, Kasper S, Ruhrmann B, Möller HJ. Hypericum in the treatment of seasonal affective disorders. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S29–33.

Mo 2004 *{published data only}*
Mo WY, Yu HY, Yang Y. Comparison of Saint John' Wort extract with fluoxetine in the treatment of post-stroke depression. *Chinese Journal of New Drugs* 2004;**13**:942–4.

Osterheider 1992 *{published and unpublished data}*
Osterheider M, Schmidtke A, Beckmann H. Behandlung depressiver Syndrome mit Hypericum (Johanniskraut - eine placebokontrollierte Doppelblindstudie. *Fortschritte Neurologie Psychiatrie* 1992; **60**(suppl.2):210–1.

Panijel 1985 *{published data only}*
Panijel J. Die Behandlung mittelschwerer Angstzustände. *Therapiewoche* 1985;**41**:4659–68.

Quandt 1993 *{published and unpublished data}*
Quandt J, Schmidt U, Schenk N. Ambulante Behandlung leichter und mittelschwerer depressiver Verstimmungen. *Der Allgemeinarzt* 1993;**15**(2):97–102.

Reh 1992 *{published and unpublished data}*
Reh C, Laux P, Schenk N. Hypericum-Extrakt bei Depressionen - eine wirksame Alternative. *Therapiewoche* 1992;**42**:1576–81.

Schlich 1987 *{published and unpublished data}*
Schlich D, Braukmann F, Schenk N. Behandlung depressiver Zustände mit Hypericinium. *Psycho* 1987;**13**:440–7.

Schmidt 1989 *{published and unpublished data}*
Schmidt U, Schenk N, Schwarz I, Vorberg G. Zur Therapie depressiver Verstimmungen. *Psycho* 1989;**15**:665–71.

Schmidt 1993 *{published and unpublished data}*
Schmidt U. Konstantes Reaktionsvermögen unter antidepressiver Theerapie mit dem Hypericum-Präparat LI 160. *Nervenheilkunde* 1991;**10**:311–2.

* Schmidt U, Sommer H. Johanniskraut-Extrakt zur ambulanten Therapie der Depression. *Fortschritte der Medizin* 1993;**111**:339–42.

Schmidt 1993b *{published data only}*
Schmidt U, Harrer G, Kuhn U, Berger-Deinert W, Luther D. Wechselwirkungen von Hypericin-Extrakt mit Alkohol. *Nervenheilkunde* 1993;**12**:314–9.

Schulz 1993 *{published data only}*
Schulz H, Jobert M. Der Einfluß von Johanniskraut auf das Schlaf-EEG bei älteren Probandinnen. *Nervenheilkunde* 1993;**12**:323–7.

Schulz H, Jobert M. Effetcs of hypericum extract on the sleep EEG in older volunteers. *Journal of Geriatric Psychiatry and Neurology* 1994; **7**(suppl 1):S39–43.

Sindrup 2000 *{published data only}*
Sindrup SS, Madsen C, Bach FW, Gram LF, Jensen TS. [St. John's wort has no effect on pain in polyneuropathy]. *Pain* 2000;**91**:361–365.

Sommer 1994 *{published and unpublished data}*
Harrer G, Sommer H. Therapie leichter/mittelschwerer Depressionen mit Hypericum. *Münchener medizinische Wochenschrift* 1993; **135**:305–9.

Harrer G, Sommer H. Treatment of mild/moderate depressions with Hypericum. *Phytomedicine* 1994;**1**:3–8.

Sommer H. Besserung psychovegetativer Beschwerden durch Hypericum im Rahmen einer multizentrischen Dopppelblindstudie. *Nervenheilkunde* 1991;**10**:308–10.

Sommer H, Harrer G. Placebo-controlled double-blind study examining the effectiveness of an hypericum preparation in 105 mildly depressed patients. *Journal of Geriatric Psychiatry and Neurology* 1994; **7**(suppl 1):S9–11.

Sommer H, Harrer G. Placebo-kontrollierte Studie zur Wirksamkeit eine Hypericumpräparates bei 105 Patienten mit Depressionen. *Nervenheilkunde* 1993;**12**:274–7.

Spielberger 1985 *{published data only}*
Spielberger E. Johanniskraut-Präparat lindert selbst mittelschwere Depressionen. *Ärztliche Praxis* 1985;**37**:2546–7.

Staffeldt 1993 *{published data only}*
Staffeldt B, Kerb R, Brockmöller J, Ploch M, Roots I. Pharmacokinetics of hypericin and pseudohypericin after oral intake of the hypericum perforatum extract LI 160 in healthy volunteers. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S47–53.

Staffeldt B, Kerb R, Brockmöller J, Ploch M, Roots I. Pharmakokinetik von Hypericin und Pseudohypericin nach oraler Einnahme des Johanniskraut-Extrakts LI 160 bei gesunden Probanden. *Nervenheilkunde* 1993;**12**:331–8.

Steger 1985 *{published and unpublished data}*
Steger W. Depressive Verstimmungen. *Zeitschrift für Allgemeinmedizin* 1985;**61**:914–8.

Volz 2002 *{published data only}*
Volz HP, Murck H, Kasper S, Möller HJ. St John's wort extract (LI 160) in somatoform disorders: results of a placebo-controlled trial. *Psychopharmacology* 2002;**164**:294–300.

Vorbach 1994 *{published and unpublished data}*
Vorbach EU, Hübner WD, Arnoldt KH. Effectiveness and tolerance of the hypericum extract LI 160 in comparison with imipramine: randomized double-blind study in 135 outpatients. *Journal of Geriatric Psychiatry and Neurology* 1994;**7**(suppl 1):S19–23.

Vorbach EU, Hübner WD, Arnoldt KH. Wirksamkeit und Verträglichkeit des Hypericum-Extraktes LI 160 im Vergleich mit Imipramin. *Nervenheilkunde* 1993;**12**:290–6.

Warnecke 1986 *{published data only}*
Warnecke G. Beeinflussung klimakterischer Depressionen. *Zeitschrift für Allgemeinmedizin* 1986;**62**:1111–3.

Werth 1989 *{published data only}*
Werth W. Psychotonin M versus Imipramin in der Chirurgie. *Der Kassenarzt* 1989;**15**:64–8.

Wienert 1991 *{published data only}*
Wienert V, Classen R, Hiller KO. Zur Frage der Photosensibilisierung von Hypericin in einer Baldrian-Johanniskraut-Kombination - klinisch-experimentelle, plazebokontrollierte Vergleichsstudie. Lübeck-Travemünde, 3rd Phytotherapy Congress 1991.

Winkel 2000 *{published data only}*
* Winkel R, Koritsch HD, Piayda H, Trenckmann. [Phytopharmakotherapie der komorbiden depressiven Störungen bei Alkoholabhängigkeit – Wirksamkeit des Johanniskrautextraktes LI 160 bei komorbid–depressiv Alkoholerkrankten]. In: Trenckmann U editor(s). *Phytopharmaka und seelische Störungen*. Neu-Isenburg: LinguaMed, 2000:43–51.

Winkel R, Koritsch HD, Piayda H, Trenckmann U. St. John's wort extract LI 160 in depressive, alcohol addicted patients. *Phytomedicine* 2000;**7 (suppl II)**:19.

Zeller 2000 *{published data only}*
Zeller K. [Once daily administration of hypericum extract (STW 3): a convenient treatment]. In: Chrubasik S, Roufogalis BD editor(s). *Herbal medicinal products for the treatment of pain*. Lismore: Southern Cross University Press, 2000:164–8.

Additional references

Caccia 2005
Caccia S. Anti-depressant-like components of hypericum perforatum extracts: an overview of their pharmacokinetics and metabolism. *Current Drug Metabolism* 2005;**6**:531–43.

Candy 2008
M Candy B, Jones L, Williams R, Tookman A, King. Psychostimulants for depression. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD006722.pub2]

CGI 1970
National Institutes of Mental Health. CGI - Clinical Global Impressions. In: Guy W editor(s). *ECDEU Assessment Manual for Psy-*

chopharmacology. Rockville, MD: US Department of Heath, Education, and Welfare Public Health Service Alcohol, Drug Abuse, and Mental Health Administration, 1970:217–222.

Deeks 2006
Deeks JJ, Higgins JPT, Altman DG on behalf of the Cochrane Statistical Methods Group. Analysing and presenting results. In: Higgins JPT, Green S editor(s). *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 [updated September 2006]. In: The Cochrane Library, Issue 4, 2006*. Chichester: John Wiley & Sons, Ltd, 2006.

DSM-IV-TR
American Psychiatric Association. DSM-IV-TR. http://www.dsmivtr.org/index.cfm.

Egger 1997
Egger M, Davey Smith G, Schneider M, Minder CE. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**:629–34.

Ernst 1995
Ernst E. St. Johns wort, an anti-depressant? A systematic, criteria-based review. *Phytomedicine* 1995;**2**:67–71.

Ernst 1999
Ernst E. Second thoughts about safety of St John's wort. *Lancet* 1999; **354**:2014–5.

Gaster 2000
Gaster B. St John's wort for depression. A systematic review. *Archives of Internal Medicine* 2000;**160**:152–6.

Hamilton 1960
Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* 1960;**23**:56–62.

Hammerness 2003
Hammerness P, Basch E, Ulbricht C. St. John's wort: a systematic review of adverse effects and drug interactions for the consultation psychiatrist. *Psychosomatics* 2003;**44**:271–82.

Higgins 2003
Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557–60.

Higgins 2008
Higgins JPT, Altman DG on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S editor(s). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 [updated February 2008]*. The Cochrane Collaboration, 2008.

Härter 2003
Härter M, Bermejo I, Schneider F, Katz S, Gaebel W, Hegerl U, Niebling W, Berger M. Clinical practice guideline for the diagnosis and treatment of depressive disorders in primary care [Versorgungsleitlinien zur Diagnostik und Therapie depressiver Störungen in der hausärztlichen Praxis]. *Zeitschrift für ärztliche Fortbildung und Qualitätssicherung* 2003;**7 (suppl. IV)**:16–35.

ICD-10
World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Version for 2007. http://www.who.int/classifications/apps/icd/icd10online/.

ICSI 2007
Institute for Clinical Systems Improvement (ICSI). Health care guideline: major depression in adults in primary care. www.icsi.org 2007.

Jadad 1996
Jadad AR, Moore RA, Carrol D, Jenkinson C, Reynolds DJM, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Controlled Clinical Trials* 1996;**17**:1–12.

Khan 2000
Khan A, Warner HA, Brown WA. Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials. *Archives of General Psychiatry* 2000;**57**:311–7.

Kim 1999
Kim HL, Streltzer J, Goebert D. St John's wort for depression: a meta-analysis of well-defined clinical trials. *Journal of Nerval and Mental Diseases* 1999;**187**:532–9.

Kirsch 2002
Kirsch I, Moore TJ, Scoboria A, Nicholls SS. The emporor's new drugs: an analysis of antidepressant medication submitted to the U.S. Food and Drug Administration. *Prevention & Treatment* 2002; **5**:article 23.

Kirsch 2008
Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* 2008;**5**:260–8.

Knüppel 2004
Knüppel L, Linde K. Adverse effects of St. John's wort - a systematic review. *Journal of Clinical Psychiatry* 2004;**65**:1470–9.

Linde 1996
Linde K, Ramirez G, Mulrow CD, Pauls A, Weidenhammer W, Melchart D. St John's wort for depression - an overview and meta-analysis of randomized clinical trials. *BMJ* 1996;**313**:253–8.

Linde 1996b
Linde K, Worku F, Melchart D, Wiesner-Zechmeister M, Pothmann R, Weinschütz T, Melchart D. Randomized clinical trials of acupuncture for asthma - a systematic review. *Forschende Komplementärmedizin* 1996;**3**:148–55.

Linde 1997
Linde K, Clausisus N, Ramirez G, Melchart D, Eitel F, Hedges LV, Jonas WB. Are the clinical effects of homoeopathy placebo effects? A meta-analysis of placebo-controlled trials. *Lancet* 1997;**350**:834–43.

Lipsey 2000
Lipsey M, Wilson DB. *Practical meta-analysis. Applied social research methods series, Vol. 49*. London: Sage Publications, 2000.

Montgomery 1979
Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* 1979;**134**:382–9.

Murck 2002
Murck H. Atypical depression and related disorders - neurobiological basis for their treatment with St. John's wort [Die atypische Depression und verwandte Erkrankungen – neurobiologische Grundlagen für ihre Behandlung mit Johanniskraut–Extrakt]. *Wiener medizinische Wochenschrift* 2002;**152**:398–403.

Murck 2005
Murck H, Fava M, Alpert J, Nierenberg AA, Mischoulon D, Otto MW, Zajecka J, Mannel M, Rosenbaum JF. Hypericum extract in patients with MDD and reversed vegetative signs: re-analysis from data of a double-blind, randomized trial of hypericum extract, fluoxetine, and placebo. *International Journal of Neuropsychopharmacology* 2005;**8**:215–21.

Nahrstedt 1997
Nahrstedt A, Butterweck V. Biologically active and other chemical constitutents of the herb hypericum perforatum L. *Pharmacopsychiatry* 1997;**30**(suppl 2):129–34.

Nieuwstraten 2006
Nieuwstraten C, Labiris NR, Holbrock A. Systematic overview of drug interactions with antidepressant medications. *Canadian Journal of Psychiatry* 2006;**51**:300–16.

Posternak 2007
Posternak MA, Zimmerman M. Therapeutic effect of follow-up assessments on antidepressant and placebo response rates in antidepressant effiacy trials. *British Journal of Psychiatry* 2007;**190**:287–92.

Reichling 2003
Reichling J, Hostanska K, Saller R. St. John's wort (Hypericum perforatum L.) - multicompound preparations versus single substances [Johanniskraut (Hypericum perforatum L.) – Vielstoffgemische kontra phytogene Einzelstoffe]. *Forschende Komplementärmedizin und Klassische Naturheilkunde* 2003;**10 (suppl 1)**:28–32.

Smith 2004
Smith L, Ernst E, Ewings P, Myers P, Smith C. Co-ingestion of herbal medicines and warfarin. *British Journal of General Practice* 2004;**54**: 439–41.

Sterne 2001
Sterne JA, Egger M, Davey Smith G. Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001;**323**:101–5.

Thompson 1999
Thompson SD, Sharp SJ. Explaining heterogeneity in meta-analyses: a comparison of methods. *Statitics in Medicine* 1999;**18**:2693–708.

Turner 2008
Turner EH, Matthews MM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* 2008;**358**:252–60.

Ustun 2004
Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murrary CJ. Global burden of depressive disorders in the year 2000. *British Journal of Psychiatry* 2004;**184**:386–92.

Volz 1997
Volz HP. Controlled clinical trials of hypericum extract in depressed patients - an overview. *Pharmacopsychiatry* 1997;**30**(suppl 2):72–6.

von Zerssen 1996
von Zerssen D. D-S. In: AMDP, CIPS editor(s). *Rating scales for psychiatry*. Weinheim: Beltz Test, 1990:section 2.8.

Wheatley 1998
Wheatley D. Hypericum extract - potential in the treatment of depression. *CNS Drugs* 1998;**9**:431–440.

Whitten 2006
Whitten DL, Myers SP, Hawrelak JA, Wohlmuth H. The effect of St John's wort extracts on CYP3A: a systematic review of prospective clinical trials. *British Journal of Clinical Pharmacology* 2006;**62**:512–26.

Williams 2000
Williams JW jr, Mulrow CD, Chiquette E, Hitchcock NP, Aguilar C, Cornell J. A systmatic review of newer pharmacotherapies for depression in adults: evidence report summary. *Annals of Internal Medicine* 2000;**132**:743–56.

Wilson 2002
Wilson DB. SPSS, Stata, and SAS macros for performing meta-analytic analyses. http://mason.gmu.edu/~dwilsonb/ma.html. last uptated: 15-1-2002. accessed: 10-11-2003.

Wurglics 2003
Wurglics M, Schulte-Löbbert S, Dingermann T, Schubert-Zsilavecs M. Rational and traditional hypericum preparations [Rationale und traditionelle Johanniskraut–Präparate]. *Deutsche Apotheker Zeitung* 2003;**143**:1454–8.

Wurglics 2006
Wurglics M, Schubert-Zsilavecz M. Hypericum perforatum: a "modern" herbal antidepresant: pharmacokinetics of active ingredients. *Clinical Pharmacokinetics* 2006;**45**:449–68.

References to other published versions of this review

Linde 1998
Linde K, Mulrow CD. St John's wort for depression. *The Cochrane Library* 1998, Issue 4.

Linde 2005a
Linde K, Mulrow CD, Berner M, Egger M. St John's wort for depression. *Cochrane Database of Systematic Reviews* 2005, Issue 2.

Linde 2005b
Linde K, Berner M, Egger M, Culrow C. St John's wort for depression - meta-analysis of randomised controlled trials. *British Journal of Psychiatry* 2005;**186**:99–107.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Behnke 2002

Methods	Concealment: unclear Blinding: double-blind Drop-outs/withdrawals: 6 of 35 in hypericum group, 3 of 35 in fluoxetine group Jadad score: 1-1-1 IV score: 1-0-1-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 70/61 Demographics: 47 female, mean age 51 (treatment group) and 48 (control group) Diagnosis: mild to modedrate depression (ICD 10 F32.0 and 32.1) Setting: multicenter trial from Denmark Baseline: HAMD score treatment group 20.0 +/- 3.2, control group 20.7 +/- 2.9
Interventions	Treatment: Hypericum extract (Calmigen) 2x1 coated tablet daily (300 mg extract) for 6 weeks Control: Fluoxetine 2x20 mg for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: von Zerssen Depression Scale, Clinical Global Impression Index (CGI), global assessment
Notes	Contract research organization which performed the study contacted for additional information but no answer was received

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Bjerkenstedt 2005

Methods	Concealment: consecutively numbered pharmacy Blinding: double-blind (double-dummy) Drop-outs/withdrawals: ITT analysis on 163 of 174 patients randomized Jadad score: 1-2-0 IV score: 1-1-1-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 177 /163 Demographics: 129 female, mean age 50 years Diagnosis: mild to moderate major depression (DSM-IV 296.31 or 32) Setting: 15 practices (psychiatry, neurology, GP) in Sweden Baseline: HAMD score hypericum group 24.9, fluoxetine group 23.8, placebo group 25.2

Bjerkenstedt 2005

(*Continued*)

Interventions	Treatment: Hypericum extract LI 160 (Jarsin 300) 3x1 coated table (900 mg) for 6 weeks Control 1: Fluoxetine 1x1 capsule (20 mg) for 6 weeks Control 2: Placebo for 4 weeks, then randomized to hypericum or fluoxetine
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD), Montgomery Asberg Depression Rating Scale (MADRS), Clinical Global Impression Index (CGI)
Notes	Due to the request of the ethical review board patients allocated to placebo had to be re-randomized to hypericum or fluoxetine after 4 weeks. Prioir to publication of the main reference authors provided some additional information; after publication of the full report no further unpublished information was considered necessary

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Bracher 2001

Methods	Concealment: consecutively numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: ITT analysis on 207 of 218 patients randomized Jadad score: not performed (see notes) IV score: not performed (see notes)
Participants	Number of patients included/analyzed: 218/207 Demographics: 167 female, mean age 44 years Diagnosis: mild to moderate major depression (DSM-IV 296.31 or 32) Setting: 17 practices (GP, various other) in Germany Baseline: HAMD score hypericum group 19.7, placebo group 19.7
Interventions	Treatment: Hypericum extract HYP611 (Felis 650) 1x1 coated table (650 mg) for 6 weeks Control: Placebo
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD), Montgomery Asberg Depression Rating Scale (MADRS), Clinical Global Impression Index (CGI) Patient-rated: von Zerssen Paranoid Depression Scale (PDS)
Notes	This trial was identified during the revision process. It is only published as a short report in a sponsored supplement to a nonscientific journal. The sponsor (Hexal AG, Holzkirchen, Germany) allowed a reviewer (KL) to extract detailed information from the full unpublished study report

Risk of bias

Item	Authors' judgement	Description

Bracher 2001

(*Continued*)

Allocation concealment?	Yes	A - adequate

Brenner 2000

Methods	Concealment: unclear Blinding: double-blind (double-dummy) Drop-outs/withdrawals: 7 of 15 hypericum patients, 3 of 15 sertraline patients Jadad score: 1-2-0 IV score: 1-0-1-1-1-0
Participants	Number of patients included/analyzed: 30/28 Demographics: 19 female, mean age 45 years Diagnosis: mild to moderate depressive disorder Setting: community hospital in USA Baseline: HAMD score hypericum group 21.3 +/- 3.2, sertraline group 21.7 +/- 2.7
Interventions	Treatment: Hypericum extract LI 160 3x1 tablet (900 mg extract) for 7 weeks Control: Sertraline 3x1 capsule (75 mg) for 7 weeks
Outcomes	Observation period: 7 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: von Zerssen Depression Scale (D-S)
Notes	Small study. High drop-out rate with more losses in the hypericum group. In spite of intention to treat analysis bias cannot be ruled out

Risk of bias

Item	**Authors' judgement**	**Description**
Allocation concealment?	Unclear	B - Unclear

Fava 2005

Methods	Concealment: unclear Blinding: double-blind Drop-outs/withdrawals: 18/45 in the hypericum group, 23 of 47 in the fluoxetine group and 22 of 43 in the placebo group Jadad score: 1-1-0 IV score: 1-0-1-0.5-0.5-0
Participants	Number of patients included/analyzed: 135/135 Demographics: 57% female, mean age 37 years Diagnosis: mild to moderate major depression Setting: 2 psychiatric outpatient departments in the US

Fava 2005

(*Continued*)	
	Baseline: HAMD score (17 items) 19.6 +/- 3.5 in the hypericum group, 19.9 +/- 2.9 in the fluoxetine and 19.6 +/- 3.1 in the placebo group
Interventions	Treatment: Hypericum extract LI 160 3x1 tablets (900 mg) for 12 weeks (+ 1 capsule placebo) Control 1: Fluoxetine 1x20 mg (capsule) for 12 weeks (+ 3x1 placebo tablets) Control 2: Placebo (1x1 capsule, 3x1 tablets)
Outcomes	Observation period: 12 weeks (+ 1 week placebo run-in) Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: Beck Depression Inventory (BDI)
Notes	High drop-out rate (intent to treat analysis); recruitment stopped before planned sample size was reached due to decision of the sponsor (lichtwer Pharma, Berlin). Only remission rates reported (and use instead of responder data for analyses)

Risk of bias

Item	**Authors' judgement**	**Description**
Allocation concealment?	Unclear	B - Unclear

Gastpar 2005

Methods	Concealment: numbered pharmacy Blinding: double-blind (double-dummy technique) Drop-out/withdrawals17 of 123 (hypericum), 19 of 118 (sertaline) Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-0.5
Participants	Patients included/analyzed: 241/118 (in per protocol analysis) Demographics: 74% female, mean age 19 years Diagnosis: moderate depression (F32.1 or F33.1) Setting: 18 primary care physicians in Germany Baseline: HAMD values 22.0 +/- 1.1 vs. 22.1 +/- 1.1
Interventions	Interventions: Hypericum extract STW3 1x1 tablet (612 mg extract) + 1 capsule placebo daily for 12 weeks Control: 1x1 tablet sertraline (50 mg) + 1 tablet placebo daily for 12 weeks After the 12 week-treatment phase there was an optional continuation phase
Outcomes	Observation period: 12 weeks for the main comparison + 12 weeks continuation phase Physician-rated: Hamitlon Depression Scale (HAMD, 17 items; response = HAMD score < 10 or at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: Adjective Mood Scale (BfS)
Notes	Additional information provided by first author and sponsor. The number of adverse effects in the hypericum group is much higher than in the sertraline group (189 vs. 112). For the number of patients reporting adverse effects the difference is less pronounced (74 vs. 60; 1 patient in the hypericum gorup had reported 26 adverse effects). In the hypericum group there was a

Gastpar 2005

(*Continued*)

	higher number of infections (48 vs. 26; not observed in any other trial before). The number of adverse effects which were possibly causally related with the treatment was higher in the sertraline group (12 vs. 16). Also the intensity of side effects was more pronounced in the sertraline group.

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Gastpar 2006

Methods	Concealment: numbered pharmacy Blinding: double-blind (double-dummy technique) Drop-outs/withdrawals: 6 of 131 (hypericum), 6 of 127 (citalopram), 8 of 130 (placebo) Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-1
Participants	Patients included/analyzed: 388/388 Demographics: 67% female, mean age 50 years Diagnosis: moderate depression (ICD-10 F32.1 or F33.1) Setting: 21 general practitioners and internists in Germany Baseline: HAMD score hypericum group 21.9 +/- 1.2, citalopram group 21.8 +/-1.2, placebo group 22.0 +/- 1.2
Interventions	Treatment: Hypericum extract STW3-VI 1x1 tablet (900 mg extract) for 6 weeks Control 1: Placebo for 6 weeks Control 2: Citalopram 1x1 tablet 20 mg for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamitlon Depression Scale (HAMD, 17 items; response = HAMD score < 10 or at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: Adjective Mood Scale (BfS)
Notes	Additional information provided by first author and sponsor

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Harrer 1993

Methods	Concealment: consecutively numbered pharmacy

Harrer 1993

(*Continued*)	
	Blinding: Double-blind Drop-outs/withdrawals: 7 of 51 in hypericum group, 9 of 51 in maprotiline group Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 102 /86 Demographics: 73 female, mean age 44 years in the Hypericum group and 48 years in the maprotiline group Diagnosis: single, moderately severe depressive episode (ICD 10 F32.1) (according to information from the sponsor patients met DSM-III-R criteria for major depression); Setting: 6 practices in Germany (neurology/psychiatry) Baseline: HAMD scores 20.5 +/- 3.7 in hypericum group and 21.5 +/- 3.9 in maprotiline group
Interventions	Treatment: Hypericum extract LI 160 (Jarsin 300) 3x1 coated tablet (900 mg extract) daily for 4 weeks Control: Maprotiline 3x1 coated tablet (75 mg) daily for 4 weeks
Outcomes	Observation period 4 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction or score < 10), Clinical Global Impression Index (CGI), global assessment Patient rated: Depression Scale von Zerssen D-S, global assessment
Notes	Additional information provided from sponsor (Lichtwer, Berlin, Germany)

Risk of bias

Item	**Authors' judgement**	**Description**
Allocation concealment?	Yes	A - Adequate

Harrer 1999

Methods	Concealment: unclear Blinding: double-blinding Drop-outs/withsdrawals: 8 of 77 hypericum patients and 16 of 84 fluoxetine patients Jadad score: 1-1-1 IV score: 1-0-1-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 161/149 Demographics: 129 female, mean age 69 years Diagnosis: mild to moderate depression (ICD 10 F32.0 or F32.1) Setting: 17 practices in Germany Baseline: HAMD score hypericum group 16.6, fluoxetine group 17.2
Interventions	Treatment: Hypericum extract LoHyp-57 2x2 coated tablets (800 mg) for 6 weeks Control: Fluoxetine 2x2 coated tablets (20 mg) for 6 weeks
Outcomes	Observation period: 6 weeks Physcian-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction or score < 11), Clinical Global Impression Index (CGI)

Harrer 1999

(*Continued*)	
	Patient-rated: Self-Rating Scale for Depression (SDS), Fragebogen Alltagsleben (German quality of life questionnaire)
Notes	Reporting of results partly insufficient

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

HDTSG 2002

Methods	Concealment: central telephone randomization Blinding: double-blind (double-dummy) Drop-outs/withdrawals: 31 of 113 in hypericum group, 32 of 116 in placebo group, and 32 of 111 in sertraline group Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 340/340 Demographics: 66% female, mean age 43 years Diagnosis: major depression (DSM-IV) Setting: 12 academic and community psychiatric research clinics in the US Baseline: HAMD scores 23.1 +/- 2.7 (hypericum), 22.7 +/- 2.7 (placebo), 22.5 +/- 2.5 (sertraline)
Interventions	Treatment: Hypericum LI 160 extract 900 to 1500 mg for 8 weeks Control 1: Placebo for 8 weeks Control 2: Sertraline 50 to 100 mg for 8 weeks
Outcomes	Observation period: 1 weeks run-in, 8 weeks treatment, 18 weeks follow-up for responders Physician-rated: Hamilton Depression Scale (HAMD, response = at least 50% reduction or score < 10), Clinical Global Impression Index (CGI), Global Assessment of Functioning (GAF) Patient-rated: Beck Depression Inventory (BDI), Sheehan Disability Scale (SDS)
Notes	Trial funded by public institution (NIH). Large proportion of patients with chronic depression. Some unblinding detected (sole trial which reported a check of blinding). Authors provided additional data

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Hänsgen 1996

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: 3 of 54 in hypericum group, 4 of 54 in placebo group Jadad score: 2-2-1 IV score: 1-1-1-1-1-0.5
Participants	Number of patients included/analyzed: 108/101 Demographics: 66 female, mean age 52 years Diagnosis: major depression (DSM-III-R, HAMD score > 15) Setting: 17 practices in Germany (neurologists/psychiatrist, general practitioners) Baseline: HAMD score 21.8 +/- 2.8 (hypericum), 20.4 +/- 3.4 (placebo)
Interventions	Treatment: Hypericum extract LI160 (Jarsin 300) 3x1 coated tablet daily (900 mg extract) for 4 weeks Control: Placebo For further 2 weeks both groups received Hypericum
Outcomes	Observation period: 4 weeks Physician-rated: Hamilton Depression Scale (HAMD, 21 items; response = at least 50% reduction or score < 10), Clinical Global Impression Index (CGI) after 2 and 4 weeks Patient-rated: Depression Scale von Zerssen (D-S), complaints check list (BEB) after 2 and 4 weeks
Notes	Additional information provided by author and sponsor (Lichtwer, Berlin, Germany). This trial was first published in 1993 (in German, 1994 in English) with 72 patients and re-published with 108 patients in 1996 (without refering to the earlier publications)

Risk of bias

Item	**Authors' judgement**	**Description**
Allocation concealment?	Yes	A - Adequate

Kalb 2001

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: none Jadad score: 2-2-1 IV score: 1-1-1-1-1-1
Participants	Patients included/analyzed: 72/72 Demographics: 48 female, mean age 48 years Diagnosis: mild to moderate depression (DSM-IV 296.21/31/22/32) Setting: 11 practices (psychiatry, internal medicine, GP) in Germany Baseline: HAMD score 19.7 +/- 3.4 hpyericum group, 20.1 +/- 2.6 placebo group
Interventions	Treatment: Hypericum extract WS 5572 (Neuroplant) 3x1 coated tablet (900 mg) for 6 weeks Control: Placebo for 6 weeks
Outcomes	Observation period: 3-7 days run-in, 6 weeks treatment

Kalb 2001

(*Continued*)	
	Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: von Zerssen Depression Scale (D-S), Patient's Global Assessment Scale (GPA)
Notes	Trial with adaptive design stopped early due to significant superiority at preplanned interim analysis

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Kasper 2006

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-out/withdrawals: 12 of 123 (hypericum 600 mg), 19 of 127 (hypericum 1200 mg), 8 of 82 (placebo) Jadad score: 2-2-1 IV score: 1-1-0.5-1-1-1
Participants	Patientns included/analyzed: 332/324 Demographics: 63% female, mean age 46 years Diagnosis: mild or moderate, single or recurrent, major depressive episode (DSM IV 296.21/22, 296.31/32) Setting: 11 psychiatric and 5 GP practices in Germany Baseline: HAMD score hypericum 600 mg group 22.8 +/- 3.3, hypericum 1200 mg group 22.6 +/-3.8, placebo group 23.6 +/- 4.2
Interventions	Treatment 1: 1 tablet hypericum extract (600 mg) daily + 1 placebo tablet for 6 weeks Treatment 2: 2x1 tablet hypericum extract (total daily extract dosage 1200 mg) daily for 6 weeks Control: 2x1 tablet placebo daily for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton depression Scale (HAMD, 17 items; response = atleast 50% score reduction), Clinical Global Impression Index (CGI), Montgomery-Asberg Depression Scale (MADRS) Patient-rated: Beck Depression Inventory (BDI), qualtiy of life (SF-36)
Notes	Randomization in 3:3:2 ratio Additional information provided by sponsor

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Laakmann 1998

Methods	Concealment: unclear Blinding: double-blind Drop-outs/withdrawals: 3 in group hypericum 1, 1 in group hypericum 2, 4 in placebo group Jadad score: 2-2-1 IV score: 1-0-1-1-1-1
Participants	Patients included/analyzed: 147/147 Demographics: 117 female, mean age 49 years Diagnosis: mild to moderate depression (DSM-IV) Setting: 11 practices in Germany Baseline: HAMD score 20.9 +/- 3.1 (hypericum 1), 20.3 +/- 2.7 (hypericum 2), 21.2 +/- 3.3 (placebo)
Interventions	Treatment 1: Hypericum extract WS 5572 (5% hyperforin) 3x1 coated tablet (900 mg) for 6 weeks Treatment 2: Hypericum extract WS 5573 (0.5% hyperforin) 3x1 coated tablet (900 mg) for 6 weeks Control: Placebo for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: von Zerssen Depression Scale (D-S), global assessment
Notes	Includes a group receiving a second hypericum extract with very low hyperforin content (not included in meta-analysis as this tests an extract which was not marketed)

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Lecrubier 2002

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: 18 of 186 (hypericum), 25 of 189 placebo) Jadad score: 1-2-0 IV score: 1-0-1-1-1-0.5
Participants	Patients included/analyzed: 375/375 Demographics: 297 female, mean age 41 years Diagnosis: mild to moderate depression (DSM-IV 296.21/22/31/32) Setting: 26 psychiatric centers and practices in France Baseline: HAMD score hypericum group 21.9 +/- 1.7, placebo group 21.9 +/- 1.7
Interventions	Treatment: Hypericum extract WS 5570 3x1 tablet (900 mg) for 6 weeks Control: placebo for 6 weeks
Outcomes	Observation period: 3-7 days run-in, 6 weeks treatment Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction),

Lecrubier 2002

(*Continued*)

	Montgomery-Asberg Depression Rating Scale (MADRS), Clinical Global Impression Index (CGI) Patient-rated: Symptom Check List (SCL-58)
Notes	Trial with preplanned interim analysis with 169 patients (no significant difference) Additional information provided by sponsor (Schwabe, Karlsruhe)

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Montgomery 2000

Methods	Concealment: unclear Blinding: double-blind Drop-outs/withdrawals: unclear Jadad score: 1-2-0 IV score: 1-0-0-1-1-0.5
Participants	Patients included/analyzed: 247/? Demographics: 183 female, mean age 43 years Diagnosis: mild to moderate depression (DSM-IV 296.2x/3x) Setting: 18 GPs and psychiatric outpatient clinics in the UK Baseline: HAMD about 22 in both groups (data extrapolated from figure in poster)
Interventions	Treatment: Hypericum extract LI 160 3x1 coated tablet (900 mg) for 12 weeks Control: Placebo for 12 weeks
Outcomes	Observation period: 3 to 7 days run-in, 12 weeks treatment Physician-rated: Hamilton depression Scale (HAMD, 17 items; response = at least 50% reduction or score < 10), Montgomery-Asberg Depression Rating Scale (MADRS), Clinical Global Impression Index (CGI) Patient-rated: None mentioned
Notes	Available as congress abstract and poster handout only. Additional information from sponsor (Lichtwer, Berlin)

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Moreno 2005

Methods	Concealment: unclear Blinding: double-blind (double-dummy technique) Drop-outs/withdrawals: 19 of 72 (no details for single groups reported) Jadad score: 1-1-0 IV score: 1-0-0-0.5-0.5-0
Participants	Number of patients included/analyzed: 72/66 Demographics: 83% female, mean age 40 years Diagnosis: mild to moderate major depression (DSM-IV criteria) Setting: Institute of Psychiatry, University Sao Paulo, Brasil Baseline: extrapolated from figure HAMD score hypericum group 15, placebo 17, fluoxetine, 15
Interventions	Treatment: Hypericum extract Iperisan 3x1 (900 mg) daily for 8 weeks Control 1: Placebo Control 2: Fluoxetine 20 mg/day for 8 weeks (1 capsule fluoxetine and 2 capsules placebo per day)
Outcomes	Observation period: 1-week prerandomzationw ash-out, 8 weeks post randomization Physician-rated: Hamilton Depression Scale (HAMD, 21 items, response = at least 50% redution, remission < 8 on the 17 item scale), Montgomery-Asberg Depression rating Scale (MADRS), Clinical Global Impression Index (CGI) Patient-rated: none
Notes	Publication misses important details. Authors inquired for additional information but no response received

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Philipp 1999

Methods	Concealment: numbered pharmacy Blinding: Double-blind Drop-outs/withdrawals: 13 of 106 (hypericum), 11 of 110 (imipramine), 9 of 47 (placebo) Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-0.5
Participants	Patients included/analyzed: 263/251 Demographics: 197 female, mean age 47 years Diagnosis: moderate depression (ICD-10 F32.1/F32.2) Setting: 18 GPs in Germany Baseline: HAMD score 22.7 +/- 4.2 (hypericum), 22.2 +/- 4.2 (imipramine), 22.7 +/- 4.0 (placebo)
Interventions	Treatment: Hypericum extract STEI 300 3x1 capsule (1050 mg) for 8 weeks Control 1: Imipramine 100 mg for 8 weeks Control 2: Placebo for 8 weeks
Outcomes	Observation period: 1 week screening, 8 weeks treatment

Philipp 1999

(*Continued*)	
	Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Hamilton Anxiety Scale (HAMA), Clinical Global Impression Scale (CGI) Patient-rated: Zung Depression Scale, Quality of life (SF-36)
Notes	Additional information from sponsor (Steiner, Berlin). Primary outcome comparison with placebo after 6 weeks, with imipramine after 8 weeks

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Schrader 1998

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: 1 of 81 (hypericum), 2 of 81 (placebo) Jadad score: 2-2-1 IV score: 1-0-0-1-1-1
Participants	Patients included/analyzed: 162/159 Demographics: 108 female, mean age 47 (hypericum) vs. 39 (placebo) Diagnosis: mild to moderate depression (F32.0/32.1) Setting: 16 private practices in Germany Baseline: HAMD values 20.1 +/- 2.8 (hypericum) vs. 18.7 +/- 3.5
Interventions	Treatment: 2x1 coated tablet (500 mg extract) Hypericum extract ZE117 daily for 6 weeks Control: Placebo
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 21 items; response = at least 50% reduction or score < 10), Clinical Global Impression Index (CGI) Patient-rated: visual analogue scale
Notes	Additional information provided by author and sponsor (Zeller AG, Romanshorn, Switzerland). Baseline HAMD score lower in placebo group. Low response rate in placebo group. In meta-analyses HAMD values were calculated for the 17-item version (data provided by author)

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Schrader 2000

Methods	Concealment: numbered pharmacy Blinding: double-blind (double-dummy) Drop-outs/withdrawals: 1 of 126 (hypericum), 1 of 113 (fluoxetine) Jadad score: 1-2-1 IV score: 1-0-1-0.5-0.5-1
Participants	Patients included/analyzed: 230/228 Demographics: 157 female, mean age 46 years Diagnosis: mild to moderate depression (ICD-10 F32.0/32.1) Setting: 7 practices (internal medicine) in Germany Baseline: HAMD values 19.6 +/- 3.1 (hypericum), 19.5 +/- 2.4 (fluoxetine)
Interventions	Treatment: Hypericum extract Ze 117 2x1 coated tablet (500 mg) for 6 weeks Control: Fluoxetine 1x1 capsule (20 mg) for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 21 items; response = at least 50% reduction or score < 10), Clinical Global Impression (CGI) Physician-rated: visual analogue scale
Notes	Additional information (particulary responder rates and 17-item HAMD scores) provided by author and sponsor (Zeller AG, Romanshorn, Switzerland)

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Shelton 2001

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals:15 of 98 (hypericum), 13 of 102 (placebo) Jadad score: 2-2-1 IV score: 1-1-1-1-1-1
Participants	Patients included/analyzed: 200/195 Demographics: 64% female, mean age 42 years Diagnosis: major depression (DSM-IV) Setting: 11 academic medical centers in the USA Baseline: mean HAMD values above 22 in both groups
Interventions	Treatment: Hypericum extract LI 160 3x1 to 4x1 tablet (900 to 1200 mg) for 8 weeks Control: placebo for 8 weeks
Outcomes	Observation period: 1 week run-in, 8 weeks treatment Phsician-rated: Hamilton Depression Scale (HAMD, 17 items, at least 50% reduction), Clinical Global Impression Index (CGI), Hamilton Anxiety Scale (HAMA)

Shelton 2001

(*Continued*)	
	Patient-rated: Beck Depression Inventory (BDI)
Notes	Average duration of current depressive episode more than 2 years. Additional information provided from authors

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Szegedi 2005

Methods	Allocation concealment: numbered pharmacy Blinding: double-blind (double dummy technique) Drop-out/wothdrawals: 17 of 125 in hypericum group, 29 of 126 in the paroxetine group Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 251/244 Demographics: 69% female, mean age 47 years Diagnosis: moderate or severe unupolar major depression (DSM IV 296.22/23, 296.32/33) Setting: 21 Psychiatric primary care centers in Germany Baseline: HAMD score 25.5.+/- 2.7 (hypericum), 25.5 +/- 2.9 (paroxetine)
Interventions	Treatment: 3x1 tablet (900 mg daily) hypericum extract WS 5570 daily for 6 weeks (in patients whose depression score was not improved by at least 20% 3x 600 mg were used after 2 weeks) Control: 1x1 tablet (20 mg) paroxetine daily for 6 weeks (in patients whose depression score was not improved by at least 20% 40 mg were used after 2 weeks) packed in capsules For each drug an indentically matched placebo was available
Outcomes	Observation period: 6 weeks (+ 16 weeks continuation phase) Physician-reated: Hamilton Depression Scale (HAMD, 17 items, at least 50% reduction), Clinical Global Impression (CGI), Montgomery-Asberg Depression Rating Scale (MADRS) Patient-rated: Beck Depression Inventory (BDI)
Notes	Additional information provided by sponsor

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Uebelhack 2004

Methods	Concelment: consecutively numbered pharmacy Blinding: double-blind Drop-out/withdrawals: none Jadad score: 2-2-1 IV score: 1-0-1-1-1-1
Participants	Number of patients included/analyzed: 140/140 Demographics: 67% female, mean age 45 years Diagnosis: moderate depressive disorder (ICD-10 F32.1 or F33.1) Setting: clinical trial center in Germany Baseline: 22-.8 +/- 1.1 in hypericum group, 22.8 +/- 1.2 in placebo group
Interventions	Treatment: 1x1 tablet (900 mg) STW3-VI daily for 6 weeks Control: Placebo
Outcomes	Observation period: 6 weeks Physician-rated: Hamitlon Depression Scale (HAMD, 17 items; response = HAMD score < 10 or at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: Adjective Mood Scale (BfS)
Notes	Trial performed in a single clinical trial unit with short recruitment period and low placebo response. Additional information from sponsor

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

van Gurp 2002

Methods	Concealment: independent pharmacist Blinding: double-blind Drop-outs/withdrawals: 16 of 45 (hypericum), 17 of 45 (sertraline) Jadad score: 2-2-1 IV score: 1-1-1-0.5-0.5-0
Participants	Patients included/analyzed: 90/87 Demographics: 61% female, mean age 40 years) Diagnosis: major depression (DSM IV) Setting: 12 community based offices for family medicine in Canada Baseline: mean HAMD values 18.8 +/- 3.6 (hypericum group), 19.7 +/- 3.5 (sertraline group)
Interventions	Treatment: 3x1 to 3x2 (in case of insufficient response at 4 weeks) capsules (900-1800 mg) Hypericum extract daily for 12 weeks Control: 3x1 to 3x2 (in case of insufficient response at 4 weeks) capsules Sertraline (50 to 100 mg) daily for 12 weeks In 9 patients of each group the dosage was doubled during the trial

van Gurp 2002

(*Continued*)

Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction or score < 10) Patient-rated: Beck Depression Inventory (BDI)
Notes	Additional information provided from author. High drop-out rate

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Volz 2000

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals:2/71 (hypericum), 5/71 (placebo) Jadad score: 1-2-0 IV score: 1-1-1-1-1-0.5
Participants	Patients included/analyzed: 142/140 Demographics: 81% female, mean age 47 years Diagnosis: mild or moderate episode of a major depression (DSM-IV) Setting: 17 practices for psychiatry, neurology, internal medicine and GP in Germany Baseline: HAMD 21.0 +/- 2.0 (hypericum group), 20.7 +/- 1.9 (placebo group)
Interventions	Treatment: 2x1 capsule (500 mg) Hypericum extract D-0496 daily for 6 weeks Control: Placebo
Outcomes	Observation period: 1 week placebo run in, 6 weeks treatment Phyiscian-rated: Hamilton Depression Scale (HAMD, 21 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: von Zerssen Depression Scale (D-S)
Notes	Manufacturer/sponsor no longer existing. Author could provide only minimal additional information

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Vorbach 1997

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: 9 of 107 in hypericum group, 14 of 102 in imipraimine group Jadad score: 2-2-1, IV score: 1-1-0.5-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 209/209 (186 per protocol) Demographics: 154 female, mean age 49 years Diagnosis: severe episode of a major depression Setting: 20 psychiatric practices Baseline: HAMD score 25.3 +/- 4.7 (hypericum group), 26.1 +/- 4.8 (imipramine group)
Interventions	Treatment: 3x2 coated tablets (1800 mg extract) Hypericum extract LI 160 daily for 6 weeks Control: 3x2 coated tablets imipramine (150 mg) daily for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: Depression-Scale (D-S von Zerssen)
Notes	Additional information provided from sponsor

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Wheatley 1997

Methods	Concealment: numbered pharmacy Blinding: double-blind Drop-outs/withdrawals: 20 of 87 in hypericum group, 24 of 78 amitriptyline group Jadad score: 2-2-0 IV score: 1-1-0.5-0.5-0.5-0.5
Participants	Number of patients included/analyzed: 165/156 Demographics: 126 female, mean age 40 years Diagnosis: major depression (DSM-IV) Setting: 18 general practices and one hospital outpatient clinic in the UK Baseline: HAMD score 20.6 +/- 2.1 (hypericum group), 20.8 +/- 2.3 (amitriptyline group)
Interventions	Treatment: 3x1 coated tablet (900 mg extract) Hypericum extract LI 160 daily for 6 weeks Control: 3x1 coated tablet amitriptyline (75 mg) daily for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; response = at least 50% reduction or score < 10), Montgomery-Asberg Reating Scale for Depression (MADRS), Clinical Global Impression Index (CGI)

Wheatley 1997

(*Continued*)

Notes	Some additional information provided from sponsor. HAMD mean values extrapolated from figure

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Witte 1995

Methods	Concealment: numbered pharmacy Blidning: double-blind Drop-outs/withdrawals: 1 of 48 in hypericum group, 1 of 49 in placebo group Jadad score: 1-2-1 IV score: 1-0-0.5-1-1-0.5
Participants	Number of patients included/analyzed: 97/95 Demographics: 64 female, mean age 43 years Diagnosis: depression (ICD-10 F32.1) Setting: 5 general practitioners in Germany Baseline: HAMD score 24.6 +/5.4 (hypericum group), 22.7 +/- 4.4 (placebo group)
Interventions	Treatment: 2x1 capsules (200 to 240 mg) Hypericum extract (Psychotonin forte) daily for 6 weeks Control: placebo
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 21 items; response = at least 50% reduction or score < 10), Clinical Global Impression Index (CGI) Patient-rated: Depression Scale (D-S von Zerssen), State Trait Anxiety Inventory (STAI)
Notes	Additional information provided by author

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	A - Adequate

Woelk 2000

Methods	Concealment: unclear Blinding: Double-blind (double dummy) Drop-outs/withdrawals: 15/157 (hypericum), 32 of 167 (imipramine) Jadad score: 2-2-0 IV score: 1-0-1-0.5-0.5-0

Woelk 2000

(*Continued*)

Participants	Patients included/analyzed: 324/324 Demographics: 71% female, mean age 46 years Diagnosis: mild to moderate depression (ICD10 F32.0 or 1, F33.0 or 1) Setting: 40 practices for psychiatry, internal medicine, GP in Germany Baseline: HAMD 22.4 +/-3.4 (hypericum), 22.1 +/- 2.9 (imipramine)
Interventions	Treatment: 2x1 coated tablet (500 mg extract) Hypericum extract ZE 117 daily for 6 weeks Control: Imipramine 150 mg daily for 6 weeks
Outcomes	Observation period: 6 weeks Physician-rated: Hamilton Depression Scale (HAMD, 17 items; at least 50% reduction), Clinical Global Impression Index (CGI) Patient-rated: global assessment
Notes	Some additional information by manufacturer of the extract

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	B - Unclear

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Agrawal 1994	Not major depression (Study in patients with fatigue and minor depression)
Albertini 1986	Not depression (RCT of hypericum in homeopathic preparation for dental neuralgia)
Bendre 1980	Not depression (CCT of homeopathic preparations of hypericum and arnica in dental practice)
Bergmann 1993	Inadequate dosage of amitriptyline (30 mg daily) - RCT included in previous versions of this review
Bernhadt 1993	RCT comparing two doses of hypericum for depression
Brockmöller 1997	RCT on pharmacokinetics and photosensitivity in healthy volunteers
Czekalla 1997	Report of results on ECG analyses undertaken in the Vorbach 1997 study
Dittmer 1992	Only a minority of patients suffered from depression (RCT of Hypericum for "psychovegetative complaints")
Ditzler 1992	Placebo-controlled randomized controlled trial in depressed patients using a fixed combination of several herbs including hypericum
Gu 2001	RCT comparing 900 mg of a hypericum extract with fluoxetine in patients with depression according to the Chinese classification CCMD-2. As it is not clear whether all included patients meet criteria for major depression the study was excluded. Response rates were similar in both groups
Halama 1991	Not major depression (placebo-controlled trial included in previous versions of the review)
Harrer 1991	Not major depression (placebo-controlled trial included in previous versions of the review)

Herberg 1992	RCT comparing hypericum and a combination of Hypericum and Valerian on concentration, reaction etc. in healthy volunteers
Hoffmann 1979	Not major depression (placebo-controlled trial included in previous versions of the review)
Hottenrott 1997	Not depression (RCT of a combination of hypericum and vitamine E to enhance performance of athlets)
Häring 1996	Placebo-controlled RCT of hypericum to prevent reactive depression in 28 patients with solid tumors undergoing chemotherapy
Hübner 1993	Not major depression (placebo-controlled trial included in previous versions of the review)
Johnson 1992	RCT on neurophysiological effects of hypericum in healthy volunteers
Johnson 1993	RCT comparing neurophysiological effects of hypericum and maprotiline in healthy volunteers
Kniebel 1988	RCT comparing a combination of hypericum and valerian vs. amitriptyline in depressive patients
Kugler 1990a	RCT comparing hypericum and bromazepam in depressed patients (included in earlier versions of this review; inclusion has now been limited to commonly recommended standard antidepressants)
Kugler 1990b	RCT on pharmacodynamics in depressed patients (reason for exclusion: no symptomatic outcomes related to depressive symptoms)
König 1993	Not major depression (placebo-controlled trial included in previous versions of the review)
Lehrl 1993	Diagnosis not decscribed as major depression in publication. An information collected from the sponsor for a previous version of the review that patients met criteria for major depression could not be verified (placebo-controlled trial included in previous versions of the review)
Lenoir 1999	RCT comparing three dosages of a hypericum extract in depressive patients
Li 2005	Not major depression (placebo-controlled RCT in depressive patients undergoing coronary artery bypass grafting)
Maisenbacher 1995	Not all patients suffered from depression (RCT of hypericum for anxiety)
Martinez 1993	RCT comparing light therapy and a combination of light therapy and hypericum in patients with seasonal affective disorder
Mo 2004	Not major depression (RCT comparing an hypericum extract and fluoxetine in post-stroke depression)
Osterheider 1992	Not major depression (placebo-controlled trial included in previous versions of the review)
Panijel 1985	Not depression (RCT of a combination of hypericum and valerian for anxiety)
Quandt 1993	Not major depression (placebo-controlled trial included in previous versions of the review)
Reh 1992	Not major depression (placebo-controlled trial included in previous versions of the review)
Schlich 1987	Not major depression (placebo-controlled trial included in previous versions of the review)
Schmidt 1989	Not major depression (placebo-controlled trial included in previous versions of the review)
Schmidt 1993	Not major depression (placebo-controlled trial included in previous versions of the review)
Schmidt 1993b	RCT investigating possible interactions of hypericum and alcohol in healthy volunteers
Schulz 1993	RCT investigating the effects of hypericum on the sleep-EEG in elderly volunteers
Sindrup 2000	RCT in patients with polyneuropathy (depression not mentioned)

(*Continued*)

Study	Reason for exclusion
Sommer 1994	Not major depression (placebo-controlled trial included in previous versions of the review)
Spielberger 1985	RCT comparing two hypericum preparations for depression
Staffeldt 1993	RCT on pharmacokinetics in healthy volunteers
Steger 1985	RCT of a comination of of hypericum and valerian vs. desipramine for depression
Volz 2002	RCT hypreicum vs. placebo in patients with somatoform disorders
Vorbach 1994	Not all patients met criteria for major depression (comparison with maprotiline; included in previous versions of the review)
Warnecke 1986	Open controlled trial of hypericum vs. diazepam in women with climacteric depression; method of allocation unclear (included in earlier version of this review; now inclusion limited to explicitely randomized, double-blind trials)
Werth 1989	RCT of hypericum vs. amitriptyline for about two weeks in patients undergoing amputation to prevent reactive depression (included in earlier version of this review)
Wienert 1991	RCT on photosensitivtity after application of a combination of hypericum and valerian
Winkel 2000	Not major depression (placebo-controlled trial included in previous versions of the review)
Zeller 2000	(Nonrandomized?) comparison of different administration schedules of a hypericum extract in patients with mild to moderate depression (not restricted to major depression)

DATA AND ANALYSES

Comparison 1. Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Responder - grouped by precision - primary analysis	18	3064	Risk Ratio (M-H, Random, 95% CI)	1.48 [1.23, 1.77]
1.1 Less precise trials	9	1020	Risk Ratio (M-H, Random, 95% CI)	1.87 [1.22, 2.87]
1.2 More precise trials	9	2044	Risk Ratio (M-H, Random, 95% CI)	1.28 [1.10, 1.49]
2 Responder - according to HAMD	16	2706	Risk Ratio (M-H, Random, 95% CI)	1.51 [1.22, 1.87]
2.1 Less precise trials	8	948	Risk Ratio (M-H, Random, 95% CI)	1.94 [1.19, 3.18]
2.2 More precise trials	8	1758	Risk Ratio (M-H, Random, 95% CI)	1.28 [1.06, 1.53]
3 Responder - according to CGI (Clinical Global Impression Index at least "much improved")	13	2306	Risk Ratio (M-H, Random, 95% CI)	1.47 [1.24, 1.74]
3.1 Less precise trials	7	869	Risk Ratio (M-H, Random, 95% CI)	1.74 [1.30, 2.33]
3.2 More precise trials	6	1437	Risk Ratio (M-H, Random, 95% CI)	1.26 [1.06, 1.50]
4 Responder - grouped by extract			Risk Ratio (M-H, Random, 95% CI)	Subtotals only
4.1 extract LI 160	6	981	Risk Ratio (M-H, Random, 95% CI)	1.31 [0.92, 1.86]
4.2 extract WS 5570	2	699	Risk Ratio (M-H, Random, 95% CI)	1.57 [0.96, 2.56]
4.3 extract WS 5572	2	170	Risk Ratio (M-H, Random, 95% CI)	1.47 [1.05, 2.06]
4.4 extract STW3-VI	2	401	Risk Ratio (M-H, Random, 95% CI)	3.59 [0.41, 31.56]
4.5 other extracts	6	813	Risk Ratio (M-H, Random, 95% CI)	1.45 [1.08, 1.93]
5 Responder among studies from German-speaking countries and other studies	18	3064	Risk Ratio (M-H, Random, 95% CI)	1.48 [1.23, 1.77]
5.1 Studies from German-speaking countries	11	1770	Risk Ratio (M-H, Random, 95% CI)	1.78 [1.42, 2.25]
5.2 Studies from other countries	7	1294	Risk Ratio (M-H, Random, 95% CI)	1.07 [0.88, 1.31]
6 Remission (HAMD score < 8 or < 7)	6	1236	Odds Ratio (M-H, Random, 95% CI)	2.77 [1.80, 4.26]

Comparison 2. Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Mean HAMD (Hamilton Rating Scale for Depression) scores after therapy	17	2871	Mean Difference (IV, Random, 95% CI)	-3.04 [-4.29, -1.78]
2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 to 3 weeks of treatment	13	2299	Mean Difference (IV, Random, 95% CI)	-1.22 [-2.07, -0.37]

3 Mean HAMD (Hamilton Rating Scale for Depression) score after 4 weeks of treatment	11	1634	Mean Difference (IV, Random, 95% CI)	-1.65 [-2.78, -0.52]
4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment	15	2578	Mean Difference (IV, Random, 95% CI)	-2.97 [-4.31, -1.63]
5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment	17	2931	Mean Difference (IV, Random, 95% CI)	-3.03 [-4.67, -1.39]
6 MADRS after treatment	3	640	Mean Difference (IV, Random, 95% CI)	-3.86 [-7.30, -0.42]
7 Difference MADRS baseline - end of treatment	4	1015	Mean Difference (IV, Random, 95% CI)	-3.01 [-4.88, -1.14]
8 Mean HAMD after treatment in studies from German-speaking countries and other studies	17	2871	Mean Difference (IV, Random, 95% CI)	-3.04 [-4.29, -1.78]
8.1 Studies from German-speaking countries	11	1720	Mean Difference (IV, Random, 95% CI)	-4.29 [-5.61, -2.97]
8.2 Studies from other countries	6	1151	Mean Difference (IV, Random, 95% CI)	-0.77 [-1.74, 0.20]
9 Mean Depression Scale von Zerssen (D-S) after therapy/difference baseline - after therapy	4	411	Mean Difference (IV, Random, 95% CI)	-3.72 [-5.32, -2.12]
10 Various self-rating scales	13	2330	Std. Mean Difference (IV, Random, 95% CI)	-0.47 [-0.64, -0.30]
10.1 von Zerssen Depression Scale (D-S) after treatment	3	313	Std. Mean Difference (IV, Random, 95% CI)	-0.62 [-1.11, -0.14]
10.2 von Zerssen Depression Scale (D-S) difference baseline - after treatment	2	170	Std. Mean Difference (IV, Random, 95% CI)	-0.90 [-2.02, 0.22]
10.3 von Zerssen Adjective Mood Scale	2	401	Std. Mean Difference (IV, Random, 95% CI)	-0.61 [-0.84, -0.37]
10.4 Beck Depression Inventory	1	195	Std. Mean Difference (IV, Random, 95% CI)	-0.28 [-0.56, 0.00]
10.5 Beck Depression Inventory difference baseline - after treatment	2	553	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.71, 0.09]
10.6 Zung Self Rating Depression Scale (SDS) difference baseline - after treatment	1	146	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.73, -0.02]
10.7 Symptom Checklist (SCL-58) depression score	1	375	Std. Mean Difference (IV, Random, 95% CI)	-0.16 [-0.37, 0.04]
10.8 von Zerssen Paranoid-Depressivitäts-Skala	1	177	Std. Mean Difference (IV, Random, 95% CI)	-0.18 [-0.48, 0.11]
11 Various self-rating scales in studies from German-speaking countries and other countries	13	2330	Std. Mean Difference (IV, Random, 95% CI)	-0.47 [-0.64, -0.30]
11.1 Studies from German-speaking countries	10	1531	Std. Mean Difference (IV, Random, 95% CI)	-0.57 [-0.77, -0.37]
11.2 Studies from other countries	3	799	Std. Mean Difference (IV, Random, 95% CI)	-0.17 [-0.31, -0.04]

Comparison 3. Safety - Hypericum mono-preparations vs. placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Number of patients discontinuing treatment/dropping out for adverse effects - primary analysis	16	2784	Odds Ratio (M-H, Random, 95% CI)	0.92 [0.45, 1.88]
2 Number of patients dropping out	16	2784	Odds Ratio (M-H, Random, 95% CI)	0.87 [0.67, 1.12]
3 Number of patients reporting adverse effects	14	2496	Odds Ratio (M-H, Random, 95% CI)	0.98 [0.78, 1.23]

Comparison 4. Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Responder (intent to treat) - primary analysis	17	2810	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.93, 1.09]
1.1 vs. older antidepressants	5	1016	Risk Ratio (M-H, Random, 95% CI)	1.02 [0.90, 1.15]
1.2 vs. SSRIs	12	1794	Risk Ratio (M-H, Random, 95% CI)	1.00 [0.90, 1.12]
2 Responder (per protocol)	17	2306	Risk Ratio (M-H, Random, 95% CI)	0.96 [0.88, 1.05]
2.1 vs. older antidepressants	5	854	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.78, 1.11]
2.2 vs. SSRIs	12	1452	Risk Ratio (M-H, Random, 95% CI)	0.97 [0.87, 1.08]
3 Responders according to CGI (Clinical Global Impression Index at least “much improved”)	12	2234	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.94, 1.09]
3.1 vs. older antidepressants	4	692	Risk Ratio (M-H, Random, 95% CI)	0.97 [0.87, 1.09]
3.2 vs. newer antidepressants	8	1542	Risk Ratio (M-H, Random, 95% CI)	1.03 [0.92, 1.15]
4 Responder among studies from German-speaking studies and other studies	17	2769	Risk Ratio (M-H, Random, 95% CI)	1.00 [0.93, 1.09]
4.1 Studies from German-speaking countries	9	1952	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.96, 1.13]
4.2 Studies from other countries	8	817	Risk Ratio (M-H, Random, 95% CI)	0.90 [0.76, 1.06]
5 Remission (HAMD score < 8)	4	685	Risk Ratio (M-H, Random, 95% CI)	1.24 [1.02, 1.50]
5.2 vs. SSRIs	4	685	Risk Ratio (M-H, Random, 95% CI)	1.24 [1.02, 1.50]

Comparison 5. Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Mean HAMD (Hamilton Rating Scale for Depression) after therapy	12	1889	Mean Difference (IV, Random, 95% CI)	-0.39 [-1.23, 0.45]
1.1 vs. older antidepressants	3	477	Mean Difference (IV, Random, 95% CI)	-0.06 [-1.82, 1.71]
1.2 vs. SSRIs	9	1412	Mean Difference (IV, Random, 95% CI)	-0.52 [-1.55, 0.51]
2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 or 3 weeks of treatment	9	1529	Mean Difference (IV, Random, 95% CI)	-0.12 [-1.02, 0.78]
2.1 vs. older antidepressants	3	477	Mean Difference (IV, Random, 95% CI)	-0.05 [-1.31, 1.20]
2.2 vs. SSRIs	6	1052	Mean Difference (IV, Random, 95% CI)	-0.25 [-1.50, 1.00]
3 Mean HAMD (Hamilton Rating Scale for Depression) scores after 4 weeks of treatment	9	1367	Mean Difference (IV, Random, 95% CI)	-0.34 [-1.48, 0.80]
3.1 vs. older antidepressants	3	477	Mean Difference (IV, Random, 95% CI)	0.02 [-1.11, 1.15]
3.2 vs. SSRIs	6	890	Mean Difference (IV, Random, 95% CI)	-0.69 [-2.44, 1.06]
4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment	10	1659	Mean Difference (IV, Random, 95% CI)	-0.34 [-1.24, 0.57]
4.1 vs. older antidepressants	2	391	Mean Difference (IV, Random, 95% CI)	-0.21 [-2.56, 2.14]
4.2 vs. SSRIs	8	1268	Mean Difference (IV, Random, 95% CI)	-0.38 [-1.46, 0.69]
5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment	10	1652	Mean Difference (IV, Random, 95% CI)	-0.35 [-1.23, 0.52]
5.1 vs. older antidepressants	1	210	Mean Difference (IV, Random, 95% CI)	-1.20 [-3.29, 0.89]
5.2 vs. SSRIs	9	1442	Mean Difference (IV, Random, 95% CI)	-0.25 [-1.21, 0.71]
6 MADRS after treatment	1	108	Mean Difference (IV, Random, 95% CI)	-0.90 [-4.73, 2.93]
6.2 vs. SSRIs	1	108	Mean Difference (IV, Random, 95% CI)	-0.90 [-4.73, 2.93]
7 Difference MADRS baseline - end of treatment	2	352	Mean Difference (IV, Random, 95% CI)	-2.90 [-5.10, -0.70]
7.2 vs. SSRIs	2	352	Mean Difference (IV, Random, 95% CI)	-2.90 [-5.10, -0.70]
8 Mean HAMD after treatment in studies from German-speaking countries and other studies	15	2423	Mean Difference (IV, Random, 95% CI)	-0.39 [-1.23, 0.45]
8.1 Studies from German-speaking countries	9	1888	Mean Difference (IV, Random, 95% CI)	-0.43 [-1.28, 0.41]
8.2 Studies from other countries	6	535	Mean Difference (IV, Random, 95% CI)	-0.44 [-2.67, 1.79]
9 Mean D-S (Depression Scale von Zerssen) scores after therapy	4	360	Mean Difference (IV, Random, 95% CI)	2.66 [0.83, 4.50]
9.1 vs. older antidepressants	2	272	Mean Difference (IV, Random, 95% CI)	2.81 [0.77, 4.85]
9.2 vs. SSRIs	2	88	Mean Difference (IV, Random, 95% CI)	2.04 [-2.13, 6.21]
10 Various self-rating scales	10	1570	Std. Mean Difference (IV, Random, 95% CI)	0.01 [-0.13, 0.15]
10.1 von Zerssen Depression Scale (D-S) after treatment	4	360	Std. Mean Difference (IV, Random, 95% CI)	0.28 [0.07, 0.49]
10.2 Beck Depression Inventory	1	83	Std. Mean Difference (IV, Random, 95% CI)	-0.01 [-0.44, 0.42]

10.3 Beck Depression Inventory difference baseline - after treatment	2	466	Std. Mean Difference (IV, Random, 95% CI)	-0.12 [-0.53, 0.29]
10.4 Zung Self Rating Depression Scale (SDS) difference baseline - after treatment	1	205	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.38, 0.17]
10.5 von Zerssen Adjective Mood Scale	2	456	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.24, 0.13]
11 Various self-rating scales in studies from German-speaking countries and other countries	10	1570	Std. Mean Difference (IV, Random, 95% CI)	0.01 [-0.13, 0.15]
11.1 Studies from German-speaking countries	6	1177	Std. Mean Difference (IV, Random, 95% CI)	-0.02 [-0.21, 0.18]
11.2 Studies from other countries	4	393	Std. Mean Difference (IV, Random, 95% CI)	0.10 [-0.10, 0.29]

Comparison 6. Safety - Hypericum mono-preparations vs. standard antidepressants

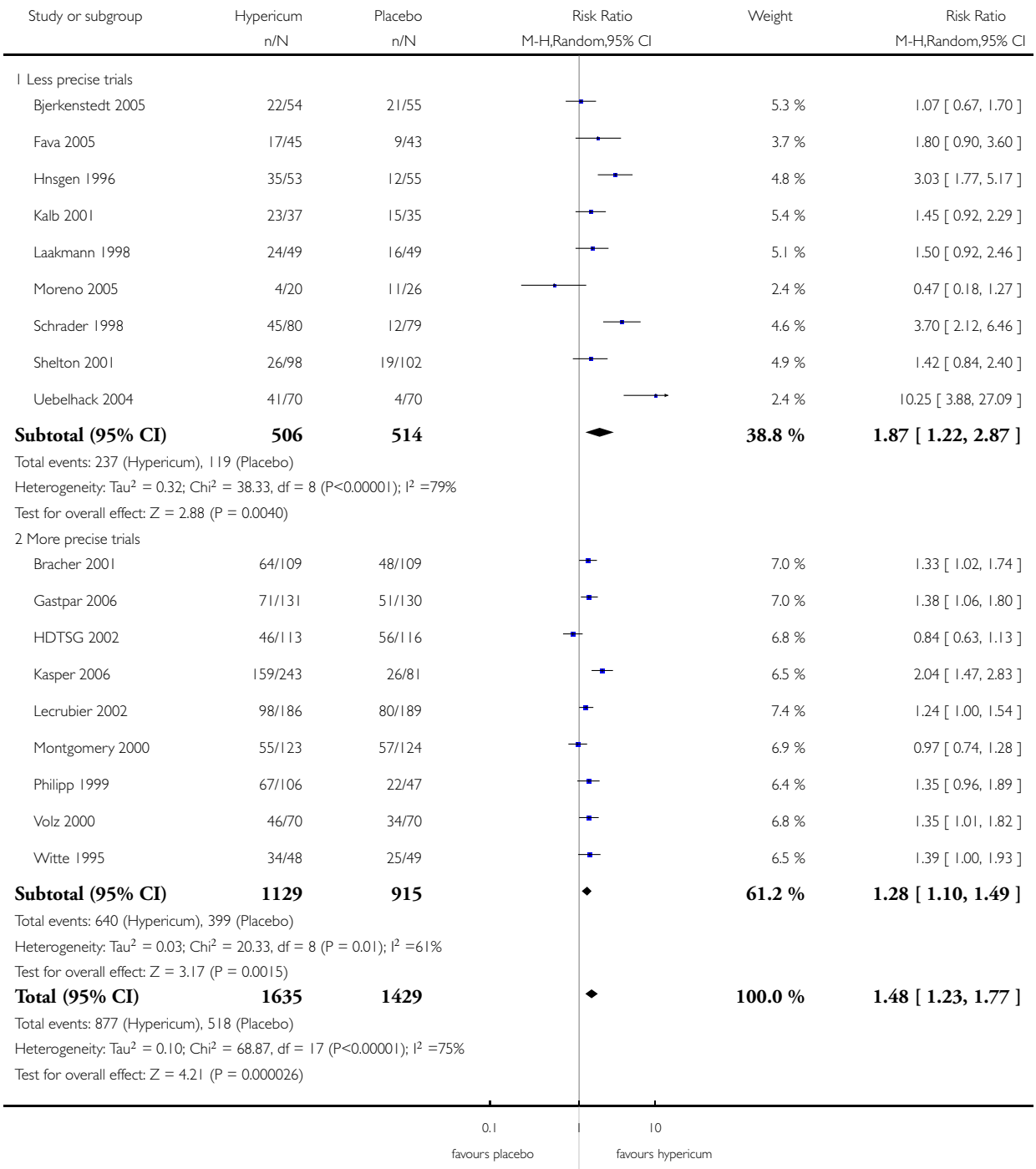
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Number of patients discontinuing treatment/dropping out due to adverse/side effects	16	2785	Odds Ratio (M-H, Random, 95% CI)	0.41 [0.29, 0.60]
1.1 vs. older antidepressants	5	1016	Odds Ratio (M-H, Random, 95% CI)	0.24 [0.13, 0.46]
1.2 vs. SSRIs	11	1769	Odds Ratio (M-H, Random, 95% CI)	0.53 [0.34, 0.83]
2 Number of patients dropping out	16	2785	Odds Ratio (M-H, Random, 95% CI)	0.77 [0.62, 0.95]
2.1 vs. older antidperessants	5	1016	Odds Ratio (M-H, Random, 95% CI)	0.67 [0.47, 0.95]
2.2 vs. SSRIs	11	1769	Odds Ratio (M-H, Random, 95% CI)	0.83 [0.63, 1.08]
3 Number of patients reporting adverse effects	14	2663	Odds Ratio (M-H, Random, 95% CI)	0.56 [0.43, 0.74]
3.1 vs. older antidepressants	5	1016	Odds Ratio (M-H, Random, 95% CI)	0.39 [0.30, 0.50]
3.2 vs. SSRIs	9	1647	Odds Ratio (M-H, Random, 95% CI)	0.70 [0.49, 1.00]

Analysis 1.1. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 1 Responder - grouped by precision - primary analysis.

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

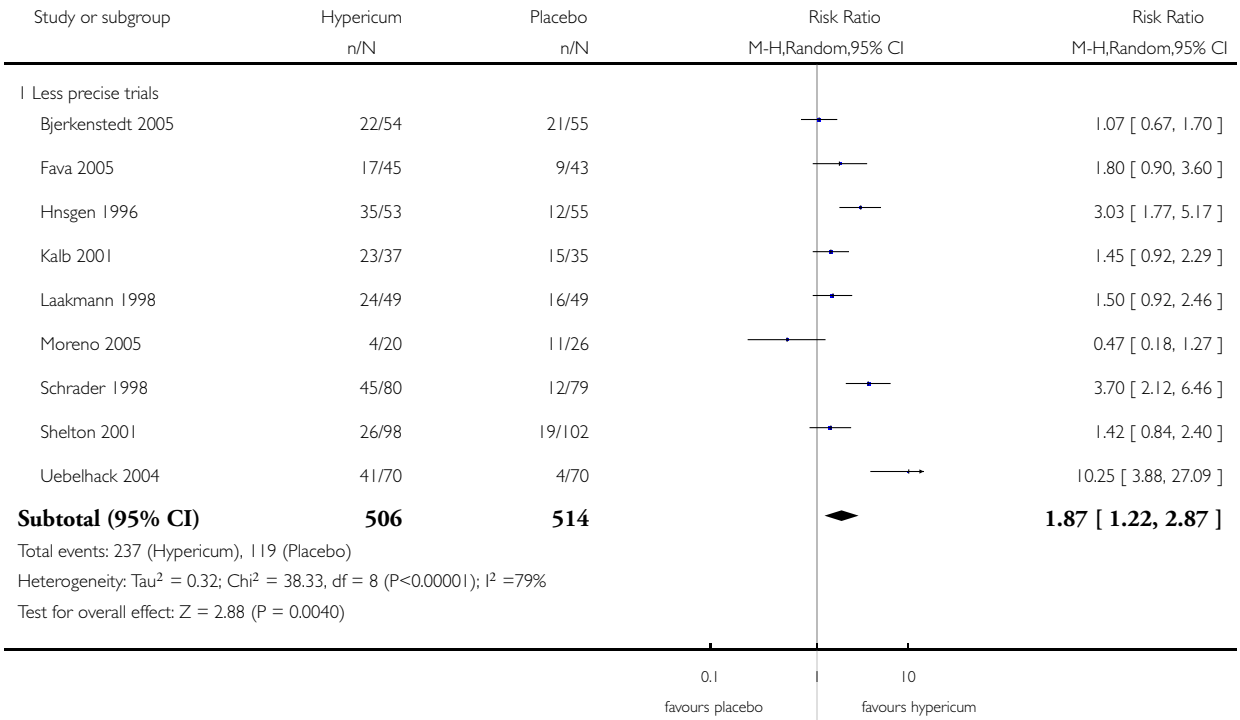
Outcome: 1 Responder - grouped by precision - primary analysis



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 1 Responder - grouped by precision - primary analysis



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 1 Responder - grouped by precision - primary analysis

Study or subgroup	Hypericum n/N	Placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 More precise trials				
Bracher 2001	64/109	48/109		1.33 [1.02, 1.74]
Gastpar 2006	71/131	51/130		1.38 [1.06, 1.80]
HDTSG 2002	46/113	56/116		0.84 [0.63, 1.13]
Kasper 2006	159/243	26/81		2.04 [1.47, 2.83]
Lecrubier 2002	98/186	80/189		1.24 [1.00, 1.54]
Montgomery 2000	55/123	57/124		0.97 [0.74, 1.28]
Philipp 1999	67/106	22/47		1.35 [0.96, 1.89]
Volz 2000	46/70	34/70		1.35 [1.01, 1.82]
Witte 1995	34/48	25/49		1.39 [1.00, 1.93]
Subtotal (95% CI)	**1129**	**915**		**1.28 [1.10, 1.49]**

Total events: 640 (Hypericum), 399 (Placebo)

Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 20.33$, df = 8 (P = 0.01); $I^2 = 61\%$

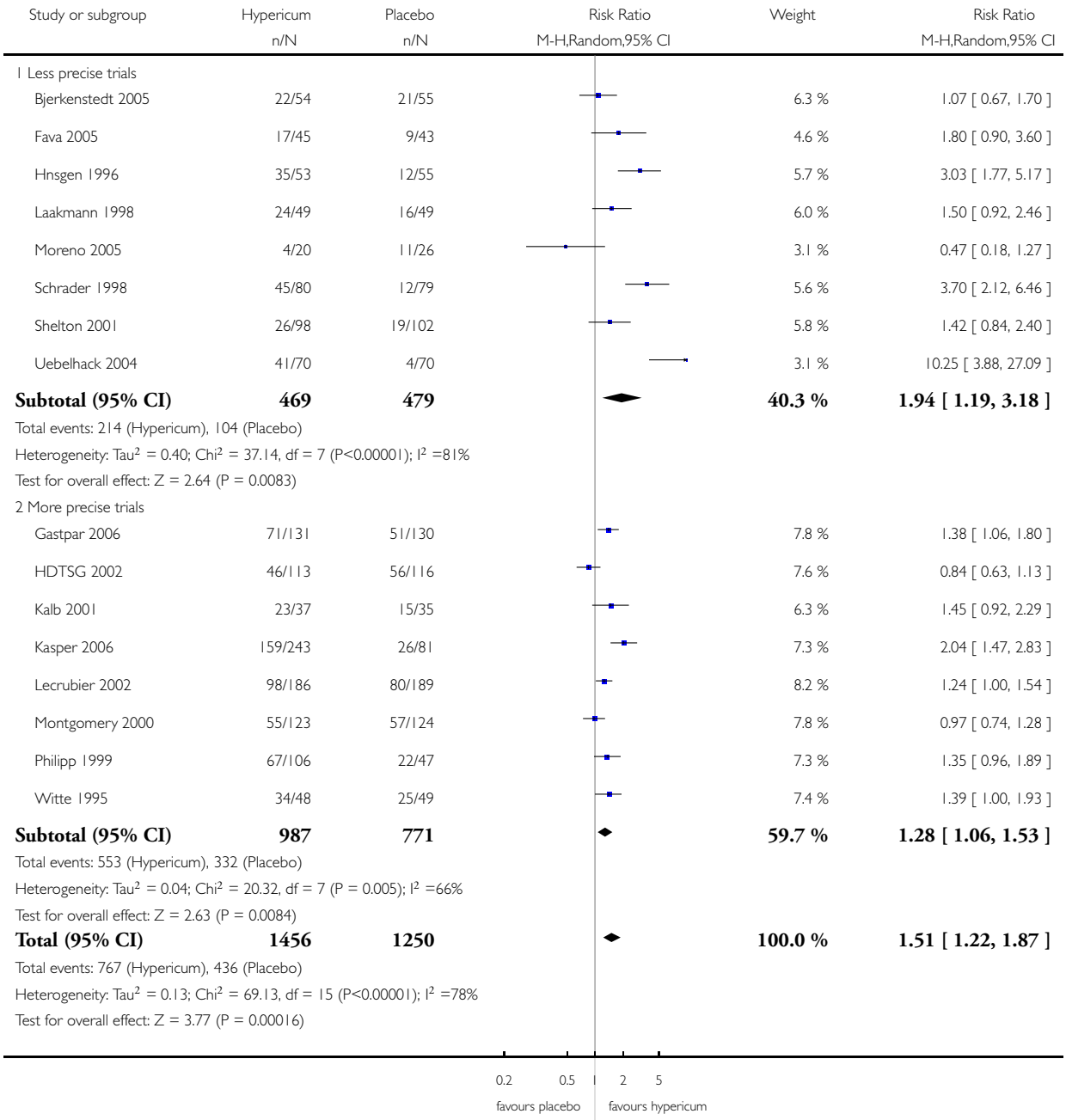
Test for overall effect: Z = 3.17 (P = 0.0015)

Analysis 1.2. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 2 Responder - according to HAMD.

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

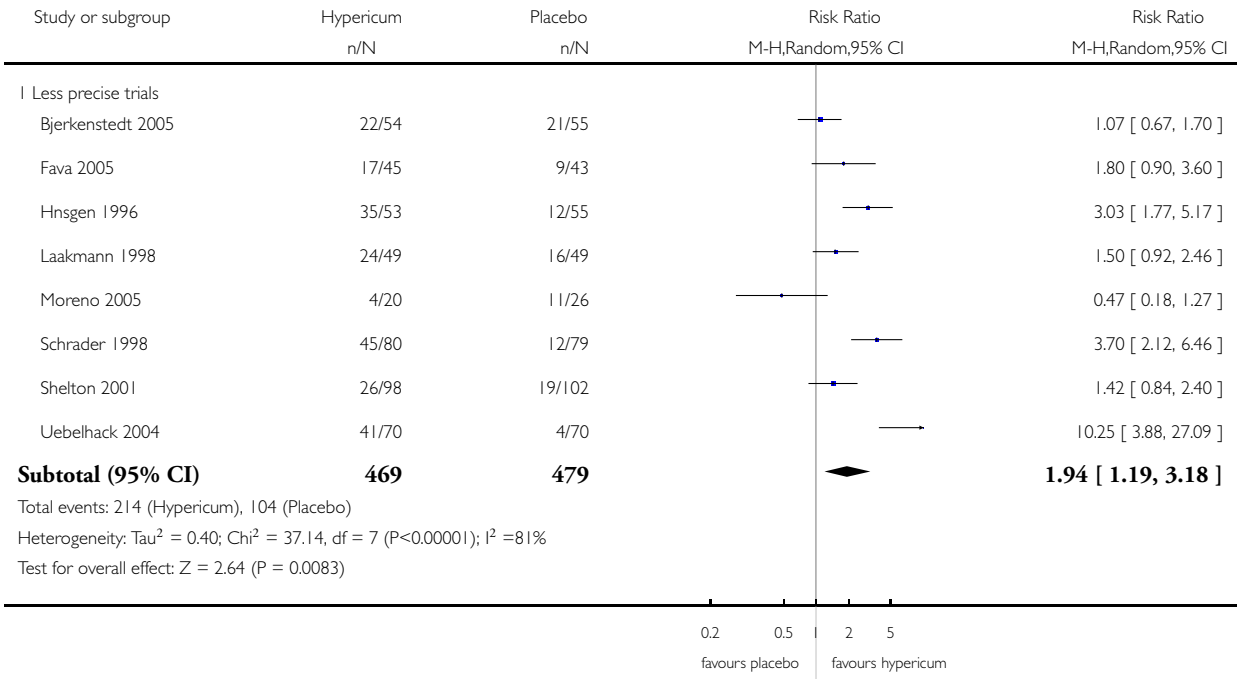
Outcome: 2 Responder - according to HAMD



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

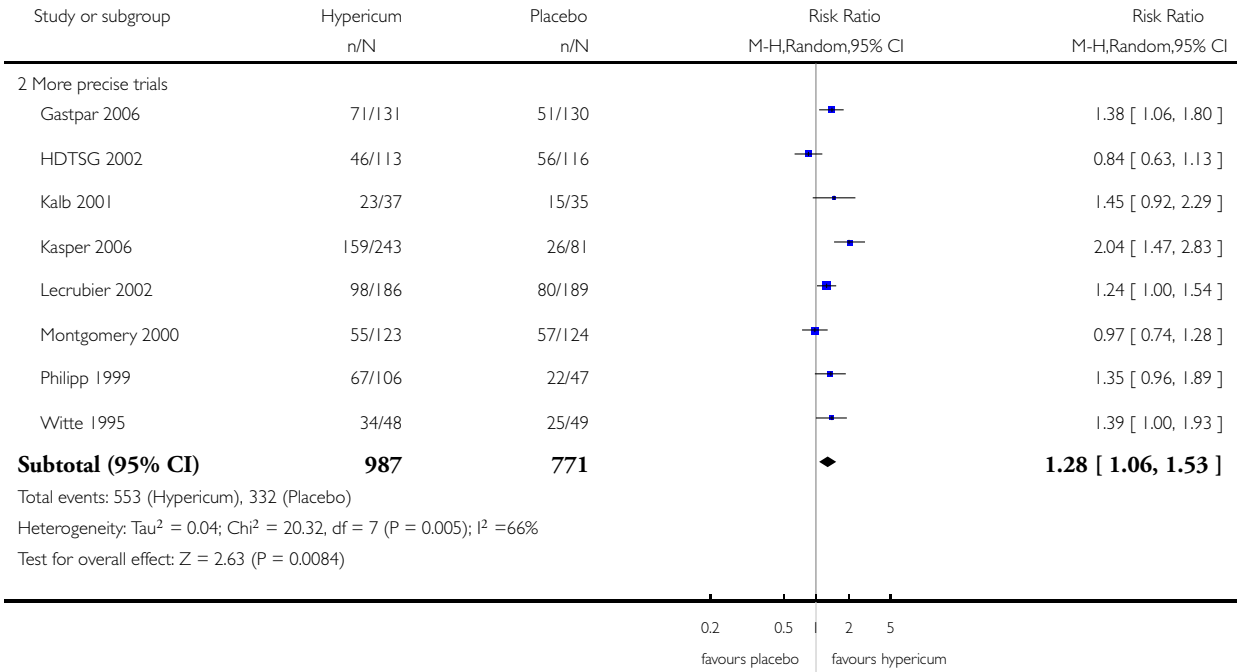
Outcome: 2 Responder - according to HAMD



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 2 Responder - according to HAMD



Analysis 1.3. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 3 Responder - according to CGI (Clinical Global Impression Index at least "much improved").

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

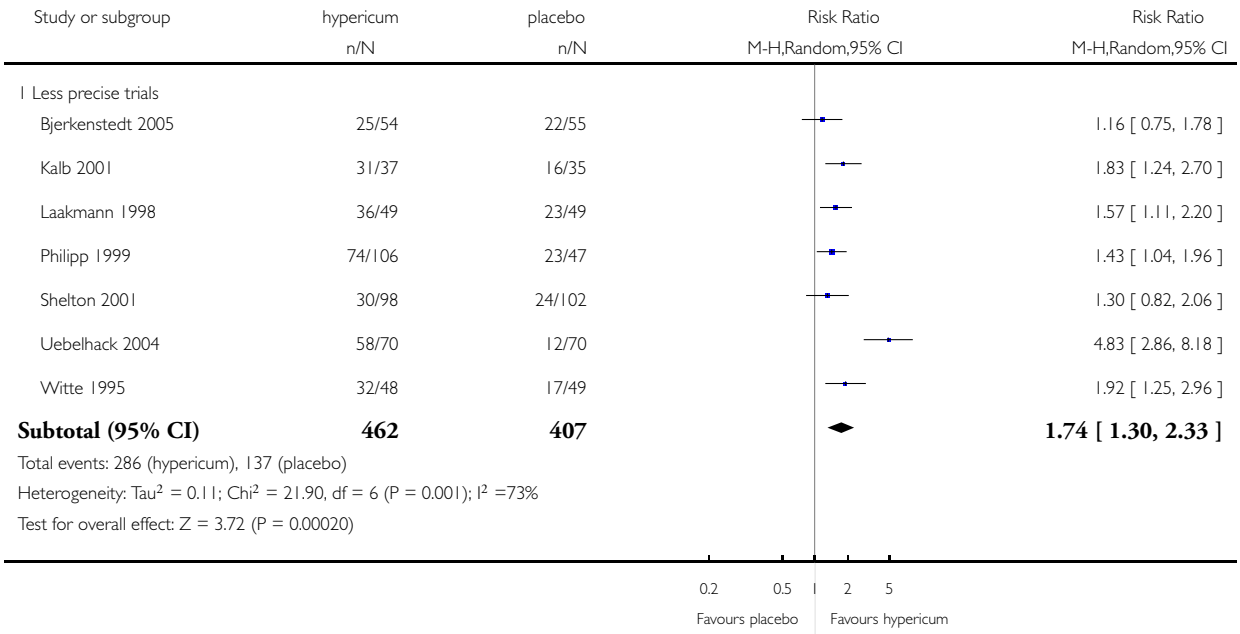
Outcome: 3 Responder - according to CGI (Clinical Global Impression Index at least "much improved")

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 Less precise trials					
Bjerkenstedt 2005	25/54	22/55		6.5 %	1.16 [0.75, 1.78]
Kalb 2001	31/37	16/35		7.0 %	1.83 [1.24, 2.70]
Laakmann 1998	36/49	23/49		7.7 %	1.57 [1.11, 2.20]
Philipp 1999	74/106	23/47		8.0 %	1.43 [1.04, 1.96]
Shelton 2001	30/98	24/102		6.1 %	1.30 [0.82, 2.06]
Uebelhack 2004	58/70	12/70		5.4 %	4.83 [2.86, 8.18]
Witte 1995	32/48	17/49		6.5 %	1.92 [1.25, 2.96]
Subtotal (95% CI)	**462**	**407**		**47.1 %**	**1.74 [1.30, 2.33]**
Total events: 286 (hypericum), 137 (placebo)					
Heterogeneity: Tau^2 = 0.11; Chi^2 = 21.90, df = 6 (P = 0.001); I^2 =73%					
Test for overall effect: Z = 3.72 (P = 0.00020)					
2 More precise trials					
Gastpar 2006	85/131	70/130		9.5 %	1.21 [0.98, 1.48]
HDTSG 2002	47/113	55/116		8.4 %	0.88 [0.66, 1.17]
Hnsgen 1996	43/53	36/55		9.2 %	1.24 [0.98, 1.56]
Kasper 2006	153/243	25/81		7.7 %	2.04 [1.45, 2.87]
Lecrubier 2002	116/186	96/189		9.8 %	1.23 [1.03, 1.47]
Volz 2000	46/70	34/70		8.3 %	1.35 [1.01, 1.82]
Subtotal (95% CI)	**796**	**641**		**52.9 %**	**1.26 [1.06, 1.50]**
Total events: 490 (hypericum), 316 (placebo)					
Heterogeneity: Tau^2 = 0.03; Chi^2 = 14.34, df = 5 (P = 0.01); I^2 =65%					
Test for overall effect: Z = 2.68 (P = 0.0075)					
Total (95% CI)	**1258**	**1048**		**100.0 %**	**1.47 [1.24, 1.74]**
Total events: 776 (hypericum), 453 (placebo)					
Heterogeneity: Tau^2 = 0.07; Chi^2 = 47.31, df = 12 (P<0.00001); I^2 =75%					
Test for overall effect: Z = 4.41 (P = 0.000010)					

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 3 Responder - according to CGI (Clinical Global Impression Index at least "much improved")



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

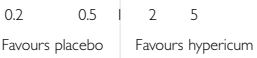
Outcome: 3 Responder - according to CGI (Clinical Global Impression Index at least "much improved")

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 More precise trials				
Gastpar 2006	85/131	70/130		1.21 [0.98, 1.48]
HDTSG 2002	47/113	55/116		0.88 [0.66, 1.17]
Hnsgen 1996	43/53	36/55		1.24 [0.98, 1.56]
Kasper 2006	153/243	25/81		2.04 [1.45, 2.87]
Lecrubier 2002	116/186	96/189		1.23 [1.03, 1.47]
Volz 2000	46/70	34/70		1.35 [1.01, 1.82]
Subtotal (95% CI)	**796**	**641**		**1.26 [1.06, 1.50]**

Total events: 490 (hypericum), 316 (placebo)

Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 14.34$, df = 5 (P = 0.01); $I^2 = 65\%$

Test for overall effect: Z = 2.68 (P = 0.0075)



Analysis 1.4. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 4 Responder - grouped by extract.

Review: St John's wort for major depression

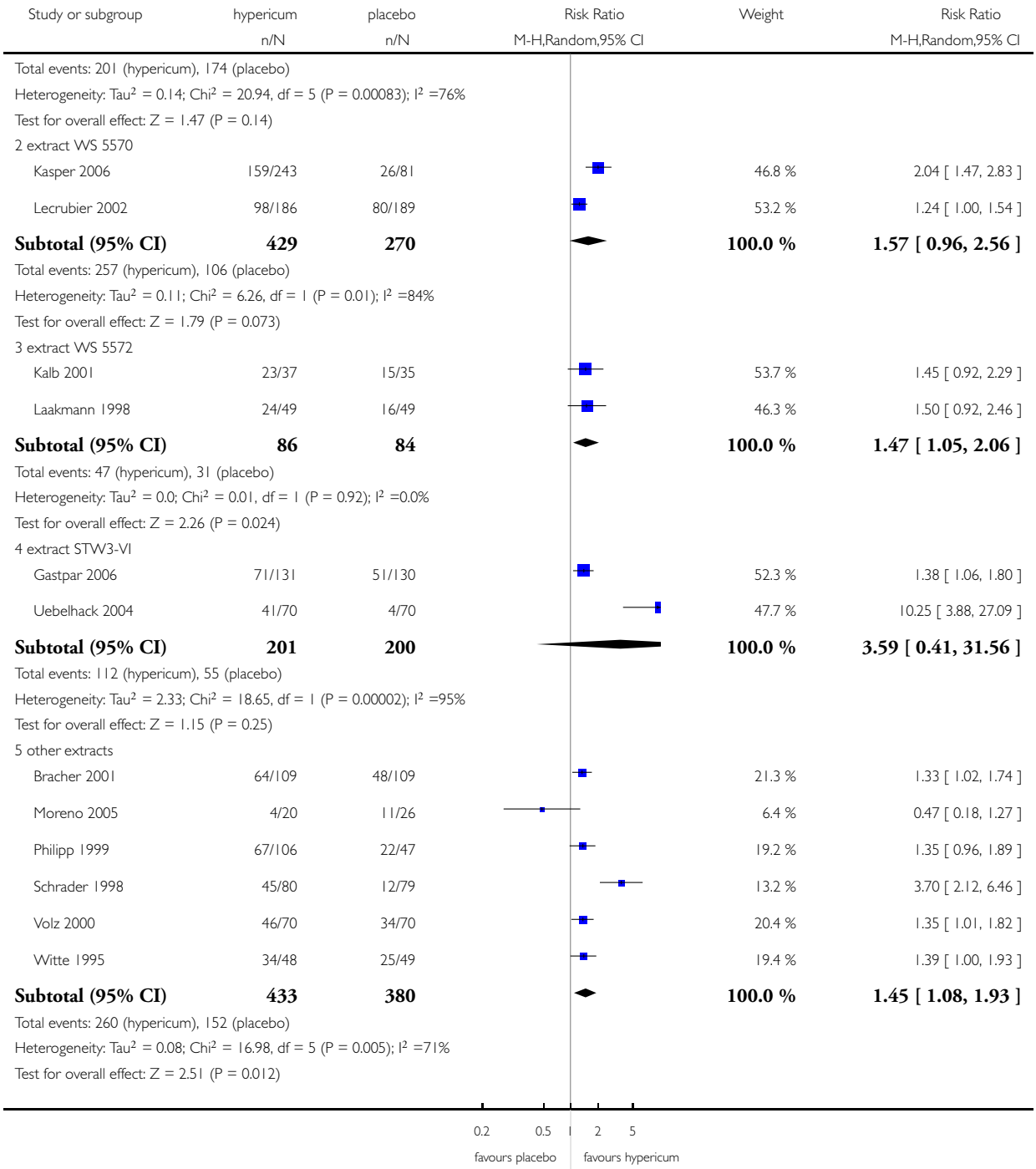
Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 4 Responder - grouped by extract

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 extract LI 160					
Bjerkenstedt 2005	22/54	21/55		16.6 %	1.07 [0.67, 1.70]
Fava 2005	17/45	9/43		12.3 %	1.80 [0.90, 3.60]
HDTSG 2002	46/113	56/116		20.1 %	0.84 [0.63, 1.13]
Hnsgen 1996	35/53	12/55		15.2 %	3.03 [1.77, 5.17]
Montgomery 2000	55/123	57/124		20.4 %	0.97 [0.74, 1.28]
Shelton 2001	26/98	19/102		15.4 %	1.42 [0.84, 2.40]
Subtotal (95% CI)	**486**	**495**		**100.0 %**	**1.31 [0.92, 1.86]**

(*Continued . . .*)

(. . . *Continued*)



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

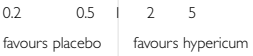
Outcome: 4 Responder - grouped by extract

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
1 extract LI 160				
Bjerkenstedt 2005	22/54	21/55		1.07 [0.67, 1.70]
Fava 2005	17/45	9/43		1.80 [0.90, 3.60]
HDTSG 2002	46/113	56/116		0.84 [0.63, 1.13]
Hnsgen 1996	35/53	12/55		3.03 [1.77, 5.17]
Montgomery 2000	55/123	57/124		0.97 [0.74, 1.28]
Shelton 2001	26/98	19/102		1.42 [0.84, 2.40]
Subtotal (95% CI)	**486**	**495**		**1.31 [0.92, 1.86]**

Total events: 201 (hypericum), 174 (placebo)

Heterogeneity: $Tau^2 = 0.14$; $Chi^2 = 20.94$, df = 5 (P = 0.00083); $I^2 = 76\%$

Test for overall effect: Z = 1.47 (P = 0.14)



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

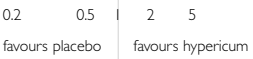
Outcome: 4 Responder - grouped by extract

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 extract WS 5570				
Kasper 2006	159/243	26/81		2.04 [1.47, 2.83]
Lecrubier 2002	98/186	80/189		1.24 [1.00, 1.54]
Subtotal (95% CI)	**429**	**270**		**1.57 [0.96, 2.56]**

Total events: 257 (hypericum), 106 (placebo)

Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 6.26$, df = 1 (P = 0.01); $I^2 = 84\%$

Test for overall effect: Z = 1.79 (P = 0.073)



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 4 Responder - grouped by extract

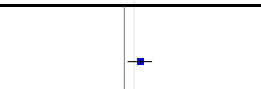
Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
3 extract WS 5572				
Kalb 2001	23/37	15/35		1.45 [0.92, 2.29]
Laakmann 1998	24/49	16/49		1.50 [0.92, 2.46]
Subtotal (95% CI)	**86**	**84**		**1.47 [1.05, 2.06]**
Total events: 47 (hypericum), 31 (placebo)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.01$, df = 1 (P = 0.92); $I^2 = 0.0\%$				
Test for overall effect: Z = 2.26 (P = 0.024)				

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 4 Responder - grouped by extract

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
4 extract STW3-VI				
Gastpar 2006	71/131	51/130		1.38 [1.06, 1.80]
Uebelhack 2004	41/70	4/70		10.25 [3.88, 27.09]
Subtotal (95% CI)	**201**	**200**		**3.59 [0.41, 31.56]**
Total events: 112 (hypericum), 55 (placebo)				
Heterogeneity: $Tau^2 = 2.33$; $Chi^2 = 18.65$, df = 1 (P = 0.00002); $I^2 = 95\%$				
Test for overall effect: Z = 1.15 (P = 0.25)				



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

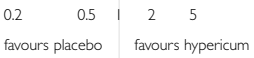
Outcome: 4 Responder - grouped by extract

Study or subgroup	hypericum n/N	placebo n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
5 other extracts				
Bracher 2001	64/109	48/109		1.33 [1.02, 1.74]
Moreno 2005	4/20	11/26		0.47 [0.18, 1.27]
Philipp 1999	67/106	22/47		1.35 [0.96, 1.89]
Schrader 1998	45/80	12/79		3.70 [2.12, 6.46]
Volz 2000	46/70	34/70		1.35 [1.01, 1.82]
Witte 1995	34/48	25/49		1.39 [1.00, 1.93]
Subtotal (95% CI)	**433**	**380**		**1.45 [1.08, 1.93]**

Total events: 260 (hypericum), 152 (placebo)

Heterogeneity: $Tau^2 = 0.08$; $Chi^2 = 16.98$, df = 5 (P = 0.005); $I^2 = 71\%$

Test for overall effect: Z = 2.51 (P = 0.012)



Analysis 1.5. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 5 Responder among studies from German-speaking countries and other studies.

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

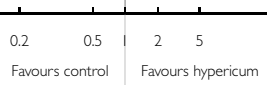
Outcome: 5 Responder among studies from German-speaking countries and other studies

Study or subgroup	Hypericum n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 Studies from German-speaking countries					
Bracher 2001	64/109	48/109		7.0 %	1.33 [1.02, 1.74]
Gastpar 2006	71/131	51/130		7.0 %	1.38 [1.06, 1.80]
Hnsgen 1996	35/53	12/55		4.8 %	3.03 [1.77, 5.17]
Kalb 2001	23/37	15/35		5.4 %	1.45 [0.92, 2.29]
Kasper 2006	159/243	26/81		6.5 %	2.04 [1.47, 2.83]
Laakmann 1998	24/49	16/49		5.1 %	1.50 [0.92, 2.46]
Philipp 1999	67/106	22/47		6.4 %	1.35 [0.96, 1.89]

(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	Hypericum n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Schrader 1998	45/80	12/79		4.6 %	3.70 [2.12, 6.46]
Uebelhack 2004	41/70	4/70		2.4 %	10.25 [3.88, 27.09]
Volz 2000	46/70	34/70		6.8 %	1.35 [1.01, 1.82]
Witte 1995	34/48	25/49		6.5 %	1.39 [1.00, 1.93]
Subtotal (95% CI)	**996**	**774**		**62.5 %**	**1.78 [1.42, 2.25]**
Total events: 609 (Hypericum), 265 (Control)					
Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 40.02$, df = 10 (P = 0.00002); I^2 =75%					
Test for overall effect: Z = 4.92 (P < 0.00001)					
2 Studies from other countries					
Bjerkenstedt 2005	22/54	21/55		5.3 %	1.07 [0.67, 1.70]
Fava 2005	17/45	9/43		3.7 %	1.80 [0.90, 3.60]
HDTSG 2002	46/113	56/116		6.8 %	0.84 [0.63, 1.13]
Lecrubier 2002	98/186	80/189		7.4 %	1.24 [1.00, 1.54]
Montgomery 2000	55/123	57/124		6.9 %	0.97 [0.74, 1.28]
Moreno 2005	4/20	11/26		2.4 %	0.47 [0.18, 1.27]
Shelton 2001	26/98	19/102		4.9 %	1.42 [0.84, 2.40]
Subtotal (95% CI)	**639**	**655**		**37.5 %**	**1.07 [0.88, 1.31]**
Total events: 268 (Hypericum), 253 (Control)					
Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 10.91$, df = 6 (P = 0.09); I^2 =45%					
Test for overall effect: Z = 0.70 (P = 0.48)					
Total (95% CI)	**1635**	**1429**		**100.0 %**	**1.48 [1.23, 1.77]**
Total events: 877 (Hypericum), 518 (Control)					
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 68.87$, df = 17 (P<0.00001); I^2 =75%					
Test for overall effect: Z = 4.21 (P = 0.000026)					

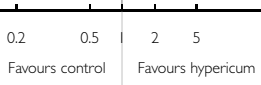


Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

Outcome: 5 Responder among studies from German-speaking countries and other studies

Study or subgroup	Hypericum n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
1 Studies from German-speaking countries				
Bracher 2001	64/109	48/109		1.33 [1.02, 1.74]
Gastpar 2006	71/131	51/130		1.38 [1.06, 1.80]
Hnsgen 1996	35/53	12/55		3.03 [1.77, 5.17]
Kalb 2001	23/37	15/35		1.45 [0.92, 2.29]
Kasper 2006	159/243	26/81		2.04 [1.47, 2.83]
Laakmann 1998	24/49	16/49		1.50 [0.92, 2.46]
Philipp 1999	67/106	22/47		1.35 [0.96, 1.89]
Schrader 1998	45/80	12/79		3.70 [2.12, 6.46]
Uebelhack 2004	41/70	4/70		10.25 [3.88, 27.09]
Volz 2000	46/70	34/70		1.35 [1.01, 1.82]
Witte 1995	34/48	25/49		1.39 [1.00, 1.93]
Subtotal (95% CI)	**996**	**774**		**1.78 [1.42, 2.25]**
Total events: 609 (Hypericum), 265 (Control)				
Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 40.02$, df = 10 (P = 0.00002); $I^2 = 75\%$				
Test for overall effect: Z = 4.92 (P < 0.00001)				



Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

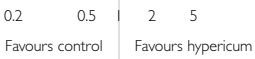
Outcome: 5 Responder among studies from German-speaking countries and other studies

Study or subgroup	Hypericum n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 Studies from other countries				
Bjerkenstedt 2005	22/54	21/55		1.07 [0.67, 1.70]
Fava 2005	17/45	9/43		1.80 [0.90, 3.60]
HDTSG 2002	46/113	56/116		0.84 [0.63, 1.13]
Lecrubier 2002	98/186	80/189		1.24 [1.00, 1.54]
Montgomery 2000	55/123	57/124		0.97 [0.74, 1.28]
Moreno 2005	4/20	11/26		0.47 [0.18, 1.27]
Shelton 2001	26/98	19/102		1.42 [0.84, 2.40]
Subtotal (95% CI)	**639**	**655**		**1.07 [0.88, 1.31]**

Total events: 268 (Hypericum), 253 (Control)

Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 10.91$, df = 6 (P = 0.09); $I^2 = 45\%$

Test for overall effect: Z = 0.70 (P = 0.48)



Analysis 1.6. Comparison 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures, Outcome 6 Remission (HAMD score < 8 or < 7).

Review: St John's wort for major depression

Comparison: 1 Hypericum mono-preparations vs. placebo A. Dichotomous measures

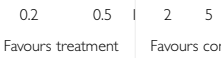
Outcome: 6 Remission (HAMD score < 8 or < 7)

Study or subgroup	Treatment n/N	Control n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Uebelhack 2004	14/70	1/70		4.1 %	17.25 [2.20, 135.23]
Bjerkenstedt 2005	13/54	4/55		10.7 %	4.04 [1.23, 13.34]
Shelton 2001	14/98	5/102		12.9 %	3.23 [1.12, 9.35]
Fava 2005	17/45	9/43		15.3 %	2.29 [0.89, 5.93]
Kasper 2006	89/243	12/81		24.6 %	3.32 [1.71, 6.47]
Lecrubier 2002	46/186	30/189		32.5 %	1.74 [1.04, 2.91]
Total (95% CI)	**696**	**540**		**100.0 %**	**2.77 [1.80, 4.26]**

Total events: 193 (Treatment), 61 (Control)

Heterogeneity: $Tau^2 = 0.08$; $Chi^2 = 7.00$, df = 5 (P = 0.22); $I^2 = 29\%$

Test for overall effect: Z = 4.65 (P < 0.00001)



Analysis 2.1. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 1 Mean HAMD (Hamilton Rating Scale for Depression) scores after therapy.

Review: St John's wort for major depression

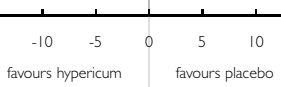
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 1 Mean HAMD (Hamilton Rating Scale for Depression) scores after therapy

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Bjerkenstedt 2005	54	15 (8.4)	55	15.5 (6.7)		5.2 %	-0.50 [-3.36, 2.36]
Bracher 2001	104	11.03 (4.96)	103	13.45 (6.36)		6.4 %	-2.42 [-3.97, -0.87]
Fava 2005	45	10.2 (6.6)	43	12.6 (6.4)		5.3 %	-2.40 [-5.12, 0.32]
Gastpar 2006	131	10.3 (6.4)	130	13 (6.9)		6.3 %	-2.70 [-4.31, -1.09]
HDTSG 2002	82	12.9 (7.1)	84	12 (7.5)		5.8 %	0.90 [-1.32, 3.12]
Hnsgen 1996	51	8.9 (4.3)	50	14.4 (5.1)		6.1 %	-5.50 [-7.34, -3.66]
Kalb 2001	37	8.9 (4.3)	35	14.4 (6.8)		5.4 %	-5.50 [-8.14, -2.86]
Kasper 2006	243	11.5 (7.7)	81	17.6 (8.8)		5.9 %	-6.10 [-8.25, -3.95]
Laakmann 1998	49	10.7 (5)	49	13.3 (6.6)		5.7 %	-2.60 [-4.92, -0.28]
Lecrubier 2002	186	12.1 (6.7)	189	13.8 (7.2)		6.5 %	-1.70 [-3.11, -0.29]
Montgomery 2000	123	12.4 (7.3)	124	12.1 (7.1)		6.2 %	0.30 [-1.50, 2.10]
Philipp 1999	100	9.2 (6.2)	46	11.9 (6.7)		5.7 %	-2.70 [-4.99, -0.41]
Schrader 1998	80	9.47 (4.82)	79	16.11 (5.7)		6.3 %	-6.64 [-8.28, -5.00]
Shelton 2001	79	15 (6.88)	87	16.11 (6.2)		6.0 %	-1.11 [-3.11, 0.89]
Uebelhack 2004	70	11.8 (4.4)	70	19.2 (3.8)		6.5 %	-7.40 [-8.76, -6.04]
Volz 2000	70	12 (5.1)	70	14.3 (5.9)		6.2 %	-2.30 [-4.13, -0.47]
Witte 1995	39	7.9 (7.5)	33	10.4 (8.1)		4.4 %	-2.50 [-6.13, 1.13]
Total (95% CI)	**1543**		**1328**			**100.0 %**	**-3.04 [-4.29, -1.78]**

Heterogeneity: $Tau^2 = 5.80$; $Chi^2 = 111.95$, df = 16 ($P<0.00001$); $I^2 = 86\%$

Test for overall effect: $Z = 4.74$ ($P < 0.00001$)



Analysis 2.2. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 to 3 weeks of treatment.

Review: St John's wort for major depression

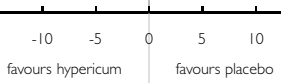
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 to 3 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Fava 2005	42	12.3 (5.9)	40	13.9 (6)		5.6 %	-1.60 [-4.18, 0.98]
Gastpar 2006	131	13.5 (5.7)	130	15.3 (6.1)		8.8 %	-1.80 [-3.23, -0.37]
HDTSG 2002	102	18.6 (5.7)	107	17.7 (5.7)		8.4 %	0.90 [-0.65, 2.45]
Hnsgen 1996	51	13.4 (3.9)	50	17.3 (4.7)		8.0 %	-3.90 [-5.59, -2.21]
Kalb 2001	37	13.9 (4.9)	35	16.1 (5.9)		5.8 %	-2.20 [-4.71, 0.31]
Kasper 2006	243	18.1 (6.5)	81	19.1 (7)		7.9 %	-1.00 [-2.73, 0.73]
Laakmann 1998	49	15.7 (4.8)	49	16.5 (5.7)		6.9 %	-0.80 [-2.89, 1.29]
Lecrubier 2002	186	15.4 (4.8)	189	16.1 (5.7)		9.9 %	-0.70 [-1.77, 0.37]
Montgomery 2000	115	17.4 (5.7)	120	17.2 (5)		9.0 %	0.20 [-1.17, 1.57]
Philipp 1999	100	16.5 (5.8)	46	17.4 (6.4)		6.6 %	-0.90 [-3.07, 1.27]
Shelton 2001	90	17.45 (5.34)	94	18.14 (4.84)		8.7 %	-0.69 [-2.16, 0.78]
Uebelhack 2004	70	16.9 (2.7)	70	20 (2.4)		10.5 %	-3.10 [-3.95, -2.25]
Witte 1995	39	19.4 (8)	33	18.8 (6.9)		4.0 %	0.60 [-2.84, 4.04]
Total (95% CI)	**1255**		**1044**			**100.0 %**	**-1.22 [-2.07, -0.37]**

Heterogeneity: $Tau^2 = 1.61$; $Chi^2 = 42.89$, df = 12 (P = 0.00002); $I^2 = 72\%$

Test for overall effect: Z = 2.80 (P = 0.0050)



Analysis 2.3. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 3 Mean HAMD (Hamilton Rating Scale for Depression) score after 4 weeks of treatment.

Review: St John's wort for major depression

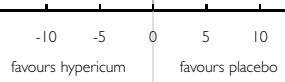
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 3 Mean HAMD (Hamilton Rating Scale for Depression) score after 4 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Bjerkenstedt 2005	54	15 (8.4)	55	15.5 (6.7)		7.4 %	-0.50 [-3.36, 2.36]
Fava 2005	35	11.5 (6.9)	35	13.9 (7.6)		6.2 %	-2.40 [-5.80, 1.00]
HDTSG 2002	97	16.9 (7.1)	99	15.9 (6.5)		10.0 %	1.00 [-0.91, 2.91]
Hnsgen 1996	51	8.9 (4.3)	50	14.4 (5.1)		10.2 %	-5.50 [-7.34, -3.66]
Kalb 2001	37	11.1 (4.5)	35	14.6 (5.9)		8.5 %	-3.50 [-5.93, -1.07]
Laakmann 1998	49	13.2 (5.5)	49	15.1 (6)		8.9 %	-1.90 [-4.18, 0.38]
Lecrubier 2002	186	13.5 (6.1)	189	14.2 (6.4)		11.9 %	-0.70 [-1.97, 0.57]
Montgomery 2000	108	13.1 (5.7)	115	13.5 (6.1)		11.0 %	-0.40 [-1.95, 1.15]
Philipp 1999	100	12.3 (5.9)	46	14.2 (7)		8.7 %	-1.90 [-4.23, 0.43]
Shelton 2001	83	16.04 (6.02)	89	16.87 (5.33)		10.6 %	-0.83 [-2.53, 0.87]
Witte 1995	39	12 (6.6)	33	14.3 (6.8)		6.8 %	-2.30 [-5.41, 0.81]
Total (95% CI)	**839**		**795**			**100.0 %**	**-1.65 [-2.78, -0.52]**

Heterogeneity: $Tau^2 = 2.38$; $Chi^2 = 32.58$, df = 10 (P = 0.00032); $I^2 = 69\%$

Test for overall effect: Z = 2.86 (P = 0.0043)



Analysis 2.4. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment.

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Bracher 2001	104	11.03 (4.96)	103	13.45 (6.36)		7.2 %	-2.42 [-3.97, -0.87]
Fava 2005	32	10.3 (5.9)	28	11.4 (5.5)		5.8 %	-1.10 [-3.99, 1.79]
Gastpar 2006	131	10.3 (6.4)	130	13 (6.9)		7.2 %	-2.70 [-4.31, -1.09]
HDTSG 2002	82	12.9 (7.1)	84	12 (7.5)		6.5 %	0.90 [-1.32, 3.12]
Kalb 2001	37	8.9 (4.3)	35	14.4 (6.8)		6.1 %	-5.50 [-8.14, -2.86]
Kasper 2006	243	11.5 (7.7)	81	17.6 (8.8)		6.6 %	-6.10 [-8.25, -3.95]
Laakmann 1998	49	10.7 (5)	49	13.3 (6.6)		6.4 %	-2.60 [-4.92, -0.28]
Lecrubier 2002	186	12.1 (6.7)	189	13.8 (7.2)		7.3 %	-1.70 [-3.11, -0.29]
Montgomery 2000	95	10.6 (6.1)	97	10.8 (6)		7.1 %	-0.20 [-1.91, 1.51]
Philipp 1999	100	9.2 (6.2)	46	11.9 (6.7)		6.5 %	-2.70 [-4.99, -0.41]
Schrader 1998	80	9.47 (4.82)	79	16.11 (5.7)		7.1 %	-6.64 [-8.28, -5.00]
Shelton 2001	79	15 (6.88)	87	16.11 (6.2)		6.8 %	-1.11 [-3.11, 0.89]
Uebelhack 2004	70	11.8 (4.4)	70	19.2 (3.8)		7.4 %	-7.40 [-8.76, -6.04]
Volz 2000	70	12 (5.1)	70	14.3 (5.9)		6.9 %	-2.30 [-4.13, -0.47]
Witte 1995	39	7.9 (7.5)	33	10.4 (8.1)		5.0 %	-2.50 [-6.13, 1.13]
Total (95% CI)	**1397**		**1181**			**100.0 %**	**-2.97 [-4.31, -1.63]**

Heterogeneity: $Tau^2 = 5.86$; $Chi^2 = 101.64$, df = 14 (P<0.00001); $I^2 = 86\%$

Test for overall effect: Z = 4.35 (P = 0.000014)

Analysis 2.5. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment.

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment

Study or subgroup	Hyericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Bjerkenstedt 2005	54	-9.9 (8.1)	55	-9.7 (7)		8.7 %	-0.20 [-3.04, 2.64]
Bracher 2001	104	-8.65 (5.27)	103	-6.27 (6.03)		10.6 %	-2.38 [-3.92, -0.84]
Fava 2005	45	-9.4 (0)	43	-7.3 (0)		0.0 %	Not estimable
Gastpar 2006	131	-11.6 (6.3)	130	-9 (6.8)		10.5 %	-2.60 [-4.19, -1.01]
HDTSG 2002	113	-8.68 (7.23)	116	-9.2 (7.22)		10.2 %	0.52 [-1.35, 2.39]
Hnsgen 1996	51	-12.1 (0)	50	-6 (0)		0.0 %	Not estimable
Kalb 2001	37	-10.8 (5)	35	-5.7 (6.4)		8.9 %	-5.10 [-7.76, -2.44]
Kasper 2006	243	-11.2 (6.8)	81	-6 (8.1)		10.0 %	-5.20 [-7.16, -3.24]
Laakmann 1998	48	-10.3 (4.6)	48	-7.9 (5.2)		10.0 %	-2.40 [-4.36, -0.44]
Lecrubier 2002	186	-9.9 (6.8)	189	-8.1 (7.1)		10.8 %	-1.80 [-3.21, -0.39]
Montgomery 2000	123	-9.2 (0)	123	-9.3 (0)		0.0 %	Not estimable
Philipp 1999	100	-13.4 (7.8)	46	-10.3 (6.1)		9.5 %	-3.10 [-5.43, -0.77]
Schrader 1998	80	-8.3 (0)	79	-1.04 (0)		0.0 %	Not estimable
Shelton 2001	79	-7.27 (0)	87	-6.66 (0)		0.0 %	Not estimable
Uebelhack 2004	70	-11.1 (4.5)	70	-3.4 (3.9)		10.8 %	-7.70 [-9.09, -6.31]
Volz 2000	70	-9 (0)	70	-6.4 (0)		0.0 %	Not estimable
Witte 1995	39	-16.7 (0)	33	-12.3 (0)		0.0 %	Not estimable
Total (95% CI)	**1573**		**1358**			**100.0 %**	**-3.03 [-4.67, -1.39]**

Heterogeneity: $Tau^2 = 5.98$; $Chi^2 = 71.43$, df = 9 ($P<0.00001$); $I^2 = 87\%$

Test for overall effect: Z = 3.62 (P = 0.00029)

Analysis 2.6. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 6 MADRS after treatment.

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 6 MADRS after treatment

Study or subgroup	Treatment N	Mean(SD)	Control N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Bjerkenstedt 2005	54	15.6 (10.4)	55	17.1 (8.7)		28.8 %	-1.50 [-5.10, 2.10]
Bracher 2001	104	11.49 (6.32)	103	14.04 (7.7)		37.2 %	-2.55 [-4.47, -0.63]
Kasper 2006	243	14.2 (9.8)	81	21.5 (10.4)		34.0 %	-7.30 [-9.88, -4.72]
Total (95% CI)	**401**		**239**			**100.0 %**	**-3.86 [-7.30, -0.42]**

Heterogeneity: $Tau^2 = 7.33$; $Chi^2 = 10.27$, df = 2 (P = 0.01); $I^2 = 81\%$

Test for overall effect: Z = 2.20 (P = 0.028)

Analysis 2.7. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 7 Difference MADRS baseline - end of treatment.

Review: St John's wort for major depression

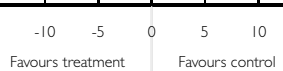
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 7 Difference MADRS baseline - end of treatment

Study or subgroup	Treatment N	Mean(SD)	Control N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Bjerkenstedt 2005	54	-9.9 (9.1)	55	-8.3 (7.9)		18.8 %	-1.60 [-4.80, 1.60]
Bracher 2001	104	-10.48 (7.12)	103	-7.76 (7.68)		28.1 %	-2.72 [-4.74, -0.70]
Kasper 2006	243	-12.5 (9.3)	81	-6.5 (10.4)		23.5 %	-6.00 [-8.55, -3.45]
Lecrubier 2002	186	-11.7 (9)	189	-9.9 (9.2)		29.7 %	-1.80 [-3.64, 0.04]
Total (95% CI)	**587**		**428**			**100.0 %**	**-3.01 [-4.88, -1.14]**

Heterogeneity: $Tau^2 = 2.18$; $Chi^2 = 7.71$, df = 3 (P = 0.05); $I^2 = 61\%$

Test for overall effect: Z = 3.16 (P = 0.0016)



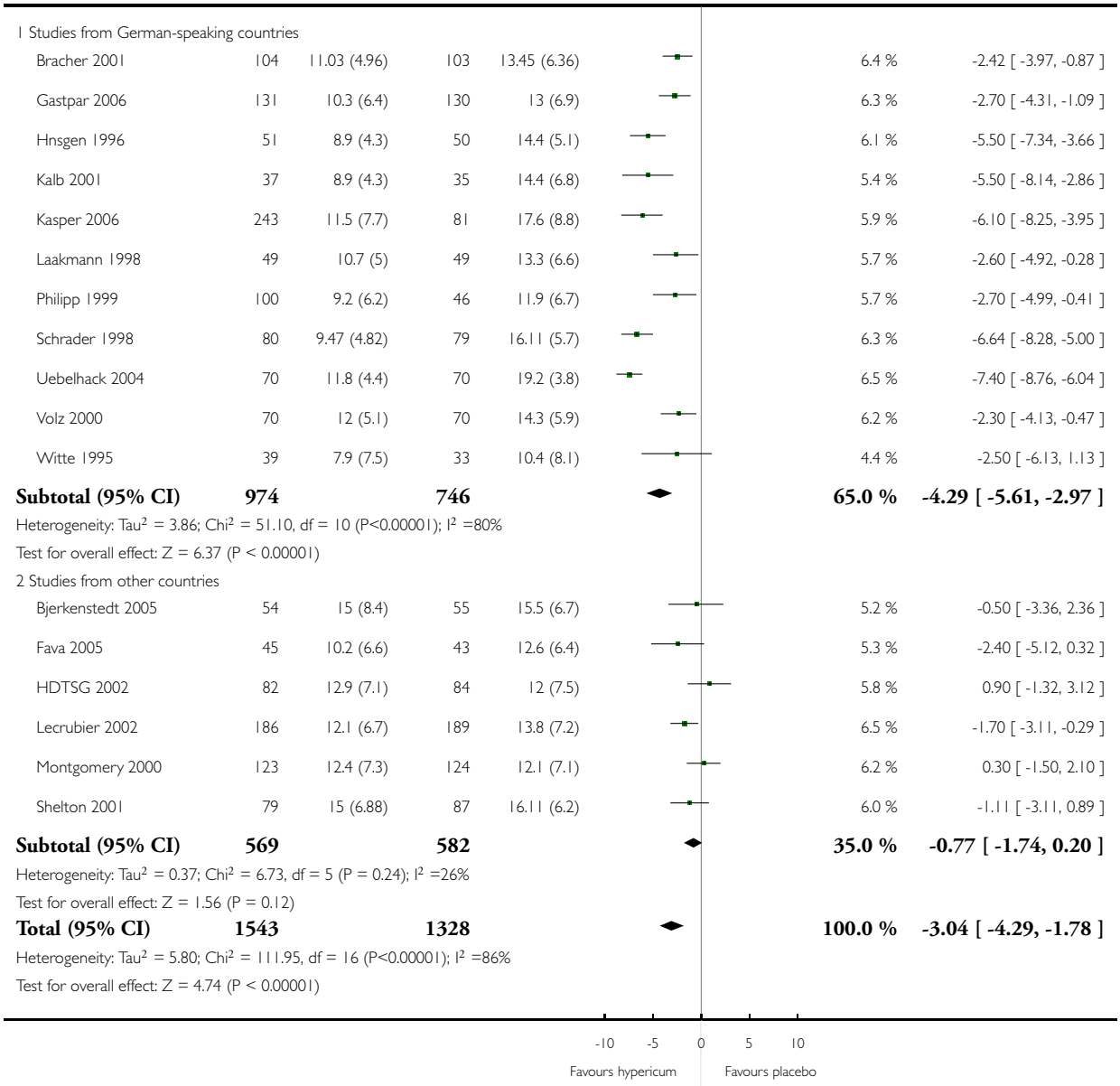
Analysis 2.8. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 8 Mean HAMD after treatment in studies from German-speaking countries and other studies.

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 8 Mean HAMD after treatment in studies from German-speaking countries and other studies

Study or subgroup	Hypericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries							
Bracher 2001	104	11.03 (4.96)	103	13.45 (6.36)		6.4 %	-2.42 [-3.97, -0.87]
Gastpar 2006	131	10.3 (6.4)	130	13 (6.9)		6.3 %	-2.70 [-4.31, -1.09]
Hnsgen 1996	51	8.9 (4.3)	50	14.4 (5.1)		6.1 %	-5.50 [-7.34, -3.66]
Kalb 2001	37	8.9 (4.3)	35	14.4 (6.8)		5.4 %	-5.50 [-8.14, -2.86]
Kasper 2006	243	11.5 (7.7)	81	17.6 (8.8)		5.9 %	-6.10 [-8.25, -3.95]
Laakmann 1998	49	10.7 (5)	49	13.3 (6.6)		5.7 %	-2.60 [-4.92, -0.28]
Philipp 1999	100	9.2 (6.2)	46	11.9 (6.7)		5.7 %	-2.70 [-4.99, -0.41]
Schrader 1998	80	9.47 (4.82)	79	16.11 (5.7)		6.3 %	-6.64 [-8.28, -5.00]
Uebelhack 2004	70	11.8 (4.4)	70	19.2 (3.8)		6.5 %	-7.40 [-8.76, -6.04]
Volz 2000	70	12 (5.1)	70	14.3 (5.9)		6.2 %	-2.30 [-4.13, -0.47]
Witte 1995	39	7.9 (7.5)	33	10.4 (8.1)		4.4 %	-2.50 [-6.13, 1.13]
Subtotal (95% CI)	**974**		**746**			**65.0 %**	**-4.29 [-5.61, -2.97]**
Heterogeneity: Tau² = 3.86; Chi² = 51.10, df = 10 (P<0.00001); I² =80%							
Test for overall effect: Z = 6.37 (P < 0.00001)							
2 Studies from other countries							
Bjerkenstedt 2005	54	15 (8.4)	55	15.5 (6.7)		5.2 %	-0.50 [-3.36, 2.36]
Fava 2005	45	10.2 (6.6)	43	12.6 (6.4)		5.3 %	-2.40 [-5.12, 0.32]
HDTSG 2002	82	12.9 (7.1)	84	12 (7.5)		5.8 %	0.90 [-1.32, 3.12]
Lecrubier 2002	186	12.1 (6.7)	189	13.8 (7.2)		6.5 %	-1.70 [-3.11, -0.29]
Montgomery 2000	123	12.4 (7.3)	124	12.1 (7.1)		6.2 %	0.30 [-1.50, 2.10]
Shelton 2001	79	15 (6.88)	87	16.11 (6.2)		6.0 %	-1.11 [-3.11, 0.89]
Subtotal (95% CI)	**569**		**582**			**35.0 %**	**-0.77 [-1.74, 0.20]**
Heterogeneity: Tau² = 0.37; Chi² = 6.73, df = 5 (P = 0.24); I² =26%							
Test for overall effect: Z = 1.56 (P = 0.12)							
Total (95% CI)	**1543**		**1328**			**100.0 %**	**-3.04 [-4.29, -1.78]**
Heterogeneity: Tau² = 5.80; Chi² = 111.95, df = 16 (P<0.00001); I² =86%							
Test for overall effect: Z = 4.74 (P < 0.00001)							



Review: St John's wort for major depression

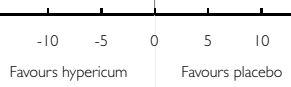
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 8 Mean HAMD after treatment in studies from German-speaking countries and other studies

Study or subgroup	Hypericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries						
Bracher 2001	104	11.03 (4.96)	103	13.45 (6.36)		-2.42 [-3.97, -0.87]
Gastpar 2006	131	10.3 (6.4)	130	13 (6.9)		-2.70 [-4.31, -1.09]
Hnsgen 1996	51	8.9 (4.3)	50	14.4 (5.1)		-5.50 [-7.34, -3.66]
Kalb 2001	37	8.9 (4.3)	35	14.4 (6.8)		-5.50 [-8.14, -2.86]
Kasper 2006	243	11.5 (7.7)	81	17.6 (8.8)		-6.10 [-8.25, -3.95]
Laakmann 1998	49	10.7 (5)	49	13.3 (6.6)		-2.60 [-4.92, -0.28]
Philipp 1999	100	9.2 (6.2)	46	11.9 (6.7)		-2.70 [-4.99, -0.41]
Schrader 1998	80	9.47 (4.82)	79	16.11 (5.7)		-6.64 [-8.28, -5.00]
Uebelhack 2004	70	11.8 (4.4)	70	19.2 (3.8)		-7.40 [-8.76, -6.04]
Volz 2000	70	12 (5.1)	70	14.3 (5.9)		-2.30 [-4.13, -0.47]
Witte 1995	39	7.9 (7.5)	33	10.4 (8.1)		-2.50 [-6.13, 1.13]
Subtotal (95% CI)	**974**		**746**			**-4.29 [-5.61, -2.97]**

Heterogeneity: $Tau^2 = 3.86$; $Chi^2 = 51.10$, df = 10 ($P<0.00001$); $I^2 = 80\%$

Test for overall effect: $Z = 6.37$ ($P < 0.00001$)



Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 8 Mean HAMD after treatment in studies from German-speaking countries and other studies

Study or subgroup	Hypericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 Studies from other countries						
Bjerkenstedt 2005	54	15 (8.4)	55	15.5 (6.7)		-0.50 [-3.36, 2.36]
Fava 2005	45	10.2 (6.6)	43	12.6 (6.4)		-2.40 [-5.12, 0.32]
HDTSG 2002	82	12.9 (7.1)	84	12 (7.5)		0.90 [-1.32, 3.12]
Lecrubier 2002	186	12.1 (6.7)	189	13.8 (7.2)		-1.70 [-3.11, -0.29]
Montgomery 2000	123	12.4 (7.3)	124	12.1 (7.1)		0.30 [-1.50, 2.10]
Shelton 2001	79	15 (6.88)	87	16.11 (6.2)		-1.11 [-3.11, 0.89]
Subtotal (95% CI)	**569**		**582**			**-0.77 [-1.74, 0.20]**

Heterogeneity: $Tau^2 = 0.37$; $Chi^2 = 6.73$, df = 5 (P = 0.24); $I^2 = 26\%$

Test for overall effect: Z = 1.56 (P = 0.12)

Analysis 2.9. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 9 Mean Depression Scale von Zerssen (D-S) after therapy/difference baseline - after therapy.

Review: St John's wort for major depression

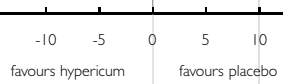
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 9 Mean Depression Scale von Zerssen (D-S) after therapy/difference baseline - after therapy

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Hnsgen 1996	51	9.3 (5)	50	14.6 (4.3)		37.9 %	-5.30 [-7.12, -3.48]
Laakmann 1998	49	-8.9 (8.2)	49	-6.4 (6.1)		22.0 %	-2.50 [-5.36, 0.36]
Volz 2000	70	10 (6)	70	12.6 (7.2)		30.9 %	-2.60 [-4.80, -0.40]
Witte 1995	39	7 (10.7)	33	10.9 (10.5)		9.3 %	-3.90 [-8.81, 1.01]
Total (95% CI)	**209**		**202**			**100.0 %**	**-3.72 [-5.32, -2.12]**

Heterogeneity: $Tau^2 = 0.89$; $Chi^2 = 4.54$, df = 3 (P = 0.21); $I^2 = 34\%$

Test for overall effect: Z = 4.56 (P < 0.00001)



Analysis 2.10. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 10 Various self-rating scales.

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
1 von Zerssen Depression Scale (D-S) after treatment							
Hnsgen 1996	51	9.3 (5)	50	14.6 (4.3)		6.6 %	-1.13 [-1.55, -0.71]
Volz 2000	70	10 (6)	70	12.6 (7.2)		7.7 %	-0.39 [-0.72, -0.06]
Witte 1995	39	7 (10.7)	33	10.9 (10.5)		6.0 %	-0.36 [-0.83, 0.10]
Subtotal (95% CI)	**160**		**153**			**20.3 %**	**-0.62 [-1.11, -0.14]**
Heterogeneity: Tau^2 = 0.14; Chi^2 = 8.51, df = 2 (P = 0.01); I^2 =77%							
Test for overall effect: Z = 2.53 (P = 0.011)							
2 von Zerssen Depression Scale (D-S) difference baseline - after treatment							
Kalb 2001	37	-15 (6.1)	35	-6 (5.9)		5.4 %	-1.48 [-2.01, -0.96]
Laakmann 1998	49	-8.9 (8.2)	49	-6.4 (6.1)		6.9 %	-0.34 [-0.74, 0.06]
Subtotal (95% CI)	**86**		**84**			**12.3 %**	**-0.90 [-2.02, 0.22]**
Heterogeneity: Tau^2 = 0.59; Chi^2 = 11.48, df = 1 (P = 0.00071); I^2 =91%							
Test for overall effect: Z = 1.58 (P = 0.11)							
3 von Zerssen Adjective Mood Scale							
Gastpar 2006	131	21.1 (11.6)	130	27.4 (12.8)		8.9 %	-0.51 [-0.76, -0.27]
Uebelhack 2004	70	20.7 (13.7)	70	31.1 (13.5)		7.6 %	-0.76 [-1.10, -0.42]
Subtotal (95% CI)	**201**		**200**			**16.4 %**	**-0.61 [-0.84, -0.37]**
Heterogeneity: Tau^2 = 0.01; Chi^2 = 1.30, df = 1 (P = 0.25); I^2 =23%							
Test for overall effect: Z = 5.09 (P < 0.00001)							
4 Beck Depression Inventory							
Shelton 2001	95	16 (9.4)	100	18.7 (9.9)		8.4 %	-0.28 [-0.56, 0.00]
Subtotal (95% CI)	**95**		**100**			**8.4 %**	**-0.28 [-0.56, 0.00]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.93 (P = 0.053)							
5 Beck Depression Inventory difference baseline - after treatment							
HDTSG 2002	113	-7.84 (9.67)	116	-6.83 (9.69)		8.7 %	-0.10 [-0.36, 0.16]
Kasper 2006	243	-8.1 (8.8)	81	-3.7 (7.9)		8.8 %	-0.51 [-0.77, -0.26]
Subtotal (95% CI)	**356**		**197**			**17.5 %**	**-0.31 [-0.71, 0.09]**
Heterogeneity: Tau^2 = 0.07; Chi^2 = 4.83, df = 1 (P = 0.03); I^2 =79%							
Test for overall effect: Z = 1.51 (P = 0.13)							
6 Zung Self Rating Depression Scale (SDS) difference baseline - after treatment							
Philipp 1999	100	-14.1 (11.3)	46	-10.2 (7.9)		7.5 %	-0.37 [-0.73, -0.02]

(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Subtotal (95% CI)	**100**		**46**			**7.5 %**	**-0.37 [-0.73, -0.02]**
Heterogeneity: not applicable							
Test for overall effect: Z = 2.09 (P = 0.037)							
7 Symptom Checklist (SCL-58) depression score							
Lecrubier 2002	186	-7.9 (8.7)	189	-6.5 (8.4)		9.4 %	-0.16 [-0.37, 0.04]
Subtotal (95% CI)	**186**		**189**			**9.4 %**	**-0.16 [-0.37, 0.04]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.58 (P = 0.11)							
8 von Zerssen Paranoid-Depressivitts-Skala							
Bracher 2001	87	-4.8 (8.8)	90	-3.1 (9.5)		8.2 %	-0.18 [-0.48, 0.11]
Subtotal (95% CI)	**87**		**90**			**8.2 %**	**-0.18 [-0.48, 0.11]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.23 (P = 0.22)							
Total (95% CI)	**1271**		**1059**			**100.0 %**	**-0.47 [-0.64, -0.30]**
Heterogeneity: Tau^2 = 0.07; Chi^2 = 46.71, df = 12 (P<0.00001); I^2 =74%							
Test for overall effect: Z = 5.34 (P < 0.00001)							

-4 -2 0 2 4

favours hypericum favours placebo

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
1 von Zerssen Depression Scale (D-S) after treatment						
Hnsgen 1996	51	9.3 (5)	50	14.6 (4.3)		-1.13 [-1.55, -0.71]
Volz 2000	70	10 (6)	70	12.6 (7.2)		-0.39 [-0.72, -0.06]
Witte 1995	39	7 (10.7)	33	10.9 (10.5)		-0.36 [-0.83, 0.10]
Subtotal (95% CI)	**160**		**153**			**-0.62 [-1.11, -0.14]**
Heterogeneity: Tau^2 = 0.14; Chi^2 = 8.51, df = 2 (P = 0.01); I^2 =77%						
Test for overall effect: Z = 2.53 (P = 0.011)						

-4 -2 0 2 4

favours hypericum favours placebo

Review: St John's wort for major depression

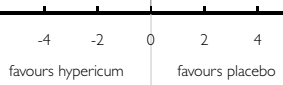
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
2 von Zerssen Depression Scale (D-S) difference baseline - after treatment						
Kalb 2001	37	-15 (6.1)	35	-6 (5.9)		-1.48 [-2.01, -0.96]
Laakmann 1998	49	-8.9 (8.2)	49	-6.4 (6.1)		-0.34 [-0.74, 0.06]
Subtotal (95% CI)	**86**		**84**			**-0.90 [-2.02, 0.22]**

Heterogeneity: $Tau^2 = 0.59$; $Chi^2 = 11.48$, df = 1 (P = 0.00071); $I^2 = 91\%$

Test for overall effect: Z = 1.58 (P = 0.11)



Review: St John's wort for major depression

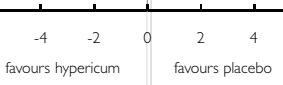
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
3 von Zerssen Adjective Mood Scale						
Gastpar 2006	131	21.1 (11.6)	130	27.4 (12.8)		-0.51 [-0.76, -0.27]
Uebelhack 2004	70	20.7 (13.7)	70	31.1 (13.5)		-0.76 [-1.10, -0.42]
Subtotal (95% CI)	**201**		**200**			**-0.61 [-0.84, -0.37]**

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 1.30$, df = 1 (P = 0.25); $I^2 = 23\%$

Test for overall effect: Z = 5.09 (P < 0.00001)



Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
4 Beck Depression Inventory						
Shelton 2001	95	16 (9.4)	100	18.7 (9.9)		-0.28 [-0.56, 0.00]
Subtotal (95% CI)	**95**		**100**			**-0.28 [-0.56, 0.00]**

Heterogeneity: not applicable

Test for overall effect: Z = 1.93 (P = 0.053)

Review: St John's wort for major depression

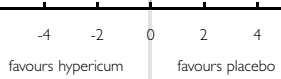
Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
5 Beck Depression Inventory difference baseline - after treatment						
HDTSG 2002	113	-7.84 (9.67)	116	-6.83 (9.69)		-0.10 [-0.36, 0.16]
Kasper 2006	243	-8.1 (8.8)	81	-3.7 (7.9)		-0.51 [-0.77, -0.26]
Subtotal (95% CI)	**356**		**197**			**-0.31 [-0.71, 0.09]**

Heterogeneity: $Tau^2 = 0.07$; $Chi^2 = 4.83$, df = 1 (P = 0.03); $I^2 = 79\%$

Test for overall effect: Z = 1.51 (P = 0.13)



Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
6 Zung Self Rating Depression Scale (SDS) difference baseline - after treatment						
Philipp 1999	100	-14.1 (11.3)	46	-10.2 (7.9)		-0.37 [-0.73, -0.02]
Subtotal (95% CI)	**100**		**46**			**-0.37 [-0.73, -0.02]**
Heterogeneity: not applicable						
Test for overall effect: Z = 2.09 (P = 0.037)						

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
7 Symptom Checklist (SCL-58) depression score						
Lecrubier 2002	186	-7.9 (8.7)	189	-6.5 (8.4)		-0.16 [-0.37, 0.04]
Subtotal (95% CI)	**186**		**189**			**-0.16 [-0.37, 0.04]**
Heterogeneity: not applicable						
Test for overall effect: Z = 1.58 (P = 0.11)						

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
8 von Zerssen Paranoid-Depressivitts-Skala						
Bracher 2001	87	-4.8 (8.8)	90	-3.1 (9.5)		-0.18 [-0.48, 0.11]
Subtotal (95% CI)	**87**		**90**			**-0.18 [-0.48, 0.11]**

Heterogeneity: not applicable

Test for overall effect: Z = 1.23 (P = 0.22)

Analysis 2.11. Comparison 2 Hypericum mono-preparations vs. placebo. B. Continuous measures, Outcome 11 Various self-rating scales in studies from German-speaking countries and other countries.

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 11 Various self-rating scales in studies from German-speaking countries and other countries

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries							
Bracher 2001	87	-4.8 (8.8)	90	-3.1 (9.5)		8.2 %	-0.18 [-0.48, 0.11]
Gastpar 2006	131	21.1 (11.6)	130	27.4 (12.8)		8.9 %	-0.51 [-0.76, -0.27]
Hnsgen 1996	51	9.3 (5)	50	14.6 (4.3)		6.6 %	-1.13 [-1.55, -0.71]
Kalb 2001	37	-15 (6.1)	35	-6 (5.9)		5.4 %	-1.48 [-2.01, -0.96]
Kasper 2006	243	-8.1 (8.8)	81	-3.7 (7.9)		8.8 %	-0.51 [-0.77, -0.26]
Laakmann 1998	49	-8.9 (8.2)	49	-6.4 (6.1)		6.9 %	-0.34 [-0.74, 0.06]
Philipp 1999	100	-14.1 (11.3)	46	-10.2 (7.9)		7.5 %	-0.37 [-0.73, -0.02]
Uebelhack 2004	70	20.7 (13.7)	70	31.1 (13.5)		7.6 %	-0.76 [-1.10, -0.42]
Volz 2000	70	10 (6)	70	12.6 (7.2)		7.7 %	-0.39 [-0.72, -0.06]
Witte 1995	39	7 (10.7)	33	10.9 (10.5)		6.0 %	-0.36 [-0.83, 0.10]
Subtotal (95% CI)	**877**		**654**			**73.5 %**	**-0.57 [-0.77, -0.37]**

Heterogeneity: $Tau^2 = 0.07$; $Chi^2 = 30.15$, df = 9 (P = 0.00041); $I^2 = 70\%$

Test for overall effect: Z = 5.61 (P < 0.00001)

(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
2 Studies from other countries							
HDTSG 2002	113	-7.84 (9.67)	116	-6.83 (9.69)		8.7 %	-0.10 [-0.36, 0.16]
Lecrubier 2002	186	-7.9 (8.7)	189	-6.5 (8.4)		9.4 %	-0.16 [-0.37, 0.04]
Shelton 2001	95	16 (9.4)	100	18.7 (9.9)		8.4 %	-0.28 [-0.56, 0.00]
Subtotal (95% CI)	**394**		**405**			**26.5 %**	**-0.17 [-0.31, -0.04]**
Heterogeneity: Tau² = 0.0; Chi² = 0.82, df = 2 (P = 0.66); I² =0.0%							
Test for overall effect: Z = 2.46 (P = 0.014)							
Total (95% CI)	**1271**		**1059**			**100.0 %**	**-0.47 [-0.64, -0.30]**
Heterogeneity: Tau² = 0.07; Chi² = 46.71, df = 12 (P<0.00001); I² =74%							
Test for overall effect: Z = 5.34 (P < 0.00001)							

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 11 Various self-rating scales in studies from German-speaking countries and other countries

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries						
Bracher 2001	87	-4.8 (8.8)	90	-3.1 (9.5)		-0.18 [-0.48, 0.11]
Gastpar 2006	131	21.1 (11.6)	130	27.4 (12.8)		-0.51 [-0.76, -0.27]
Hnsgen 1996	51	9.3 (5)	50	14.6 (4.3)		-1.13 [-1.55, -0.71]
Kalb 2001	37	-15 (6.1)	35	-6 (5.9)		-1.48 [-2.01, -0.96]
Kasper 2006	243	-8.1 (8.8)	81	-3.7 (7.9)		-0.51 [-0.77, -0.26]
Laakmann 1998	49	-8.9 (8.2)	49	-6.4 (6.1)		-0.34 [-0.74, 0.06]
Philipp 1999	100	-14.1 (11.3)	46	-10.2 (7.9)		-0.37 [-0.73, -0.02]
Uebelhack 2004	70	20.7 (13.7)	70	31.1 (13.5)		-0.76 [-1.10, -0.42]
Volz 2000	70	10 (6)	70	12.6 (7.2)		-0.39 [-0.72, -0.06]
Witte 1995	39	7 (10.7)	33	10.9 (10.5)		-0.36 [-0.83, 0.10]
Subtotal (95% CI)	**877**		**654**			**-0.57 [-0.77, -0.37]**
Heterogeneity: Tau² = 0.07; Chi² = 30.15, df = 9 (P = 0.00041); I² =70%						
Test for overall effect: Z = 5.61 (P < 0.00001)						

Review: St John's wort for major depression

Comparison: 2 Hypericum mono-preparations vs. placebo. B. Continuous measures

Outcome: 11 Various self-rating scales in studies from German-speaking countries and other countries

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
2 Studies from other countries						
HDTSG 2002	113	-7.84 (9.67)	116	-6.83 (9.69)		-0.10 [-0.36, 0.16]
Lecrubier 2002	186	-7.9 (8.7)	189	-6.5 (8.4)		-0.16 [-0.37, 0.04]
Shelton 2001	95	16 (9.4)	100	18.7 (9.9)		-0.28 [-0.56, 0.00]
Subtotal (95% CI)	**394**		**405**			**-0.17 [-0.31, -0.04]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.82$, df = 2 (P = 0.66); $I^2 = 0.0\%$

Test for overall effect: Z = 2.46 (P = 0.014)

Analysis 3.1. Comparison 3 Safety - Hypericum mono-preparations vs. placebo, Outcome 1 Number of patients discontinuing treatment/dropping out for adverse effects - primary analysis.

Review: St John's wort for major depression

Comparison: 3 Safety - Hypericum mono-preparations vs. placebo

Outcome: 1 Number of patients discontinuing treatment/dropping out for adverse effects - primary analysis

Study or subgroup	hypericum n/N	placebo n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Bjerkenstedt 2005	4/57	2/57		17.1 %	2.08 [0.36, 11.81]
Bracher 2001	0/109	0/109		0.0 %	Not estimable
Fava 2005	0/45	0/43		0.0 %	Not estimable
Gastpar 2006	4/131	6/130		31.1 %	0.65 [0.18, 2.36]
HDTSG 2002	2/113	3/116		15.8 %	0.68 [0.11, 4.14]
Hnsgen 1996	0/53	0/55		0.0 %	Not estimable
Kalb 2001	0/37	0/35		0.0 %	Not estimable
Kasper 2006	6/250	0/82		6.2 %	4.39 [0.24, 78.71]

(*Continued . . .*)

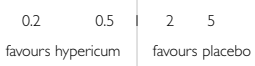
(. . . *Continued*)

Study or subgroup	hypericum n/N	placebo n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Laakmann 1998	0/49	1/49		5.0 %	0.33 [0.01, 8.22]
Lecrubier 2002	2/186	2/189		13.3 %	1.02 [0.14, 7.29]
Philipp 1999	0/106	0/47		0.0 %	Not estimable
Schrader 1998	0/80	1/79		5.0 %	0.33 [0.01, 8.10]
Shelton 2001	1/98	1/102		6.7 %	1.04 [0.06, 16.88]
Uebelhack 2004	0/70	0/70		0.0 %	Not estimable
Volz 2000	0/70	0/70		0.0 %	Not estimable
Witte 1995	0/48	0/49		0.0 %	Not estimable
Total (95% CI)	**1502**	**1282**		**100.0 %**	**0.92 [0.45, 1.88]**

Total events: 19 (hypericum), 16 (placebo)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.19$, df = 7 (P = 0.87); $I^2 = 0.0\%$

Test for overall effect: Z = 0.23 (P = 0.82)



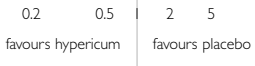
Analysis 3.2. Comparison 3 Safety - Hypericum mono-preparations vs. placebo, Outcome 2 Number of patients dropping out.

Review: St John's wort for major depression

Comparison: 3 Safety - Hypericum mono-preparations vs. placebo

Outcome: 2 Number of patients dropping out

Study or subgroup	hypericum n/N	placebo n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Bjerkenstedt 2005	9/57	5/57		4.9 %	1.95 [0.61, 6.23]
Bracher 2001	5/109	3/109		3.1 %	1.70 [0.40, 7.29]
Fava 2005	18/45	22/43		9.3 %	0.64 [0.27, 1.48]
Gastpar 2006	6/131	8/130		5.6 %	0.73 [0.25, 2.17]
HDTSG 2002	31/113	32/116		19.6 %	0.99 [0.56, 1.77]
Hnsgen 1996	2/53	4/55		2.2 %	0.50 [0.09, 2.85]
Kalb 2001	0/37	0/35		0.0 %	Not estimable
Kasper 2006	31/250	8/82		9.8 %	1.31 [0.58, 2.97]



(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	hypericum n/N	placebo n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Laakmann 1998	2/49	3/49		2.0 %	0.65 [0.10, 4.09]
Lecrubier 2002	18/186	25/189		16.0 %	0.70 [0.37, 1.34]
Philipp 1999	13/106	9/47		7.6 %	0.59 [0.23, 1.50]
Schrader 1998	1/80	2/79		1.1 %	0.49 [0.04, 5.49]
Shelton 2001	15/98	13/102		10.3 %	1.24 [0.56, 2.76]
Uebelhack 2004	0/70	0/70		0.0 %	Not estimable
Volz 2000	1/70	4/70		1.3 %	0.24 [0.03, 2.20]
Witte 1995	9/48	13/49		7.1 %	0.64 [0.24, 1.67]
Total (95% CI)	**1502**	**1282**		**100.0 %**	**0.87 [0.67, 1.12]**

Total events: 161 (hypericum), 151 (placebo)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 8.68$, df = 13 (P = 0.80); $I^2 = 0.0\%$

Test for overall effect: Z = 1.09 (P = 0.28)

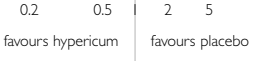
Analysis 3.3. Comparison 3 Safety - Hypericum mono-preparations vs. placebo, Outcome 3 Number of patients reporting adverse effects.

Review: St John's wort for major depression

Comparison: 3 Safety - Hypericum mono-preparations vs. placebo

Outcome: 3 Number of patients reporting adverse effects

Study or subgroup	hypericum n/N	placebo n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Bjerkenstedt 2005	20/57	18/57		7.4 %	1.17 [0.54, 2.55]
Bracher 2001	30/109	28/109		11.4 %	1.10 [0.60, 2.00]
Gastpar 2006	39/131	46/130		14.3 %	0.77 [0.46, 1.30]
HDTSG 2002	100/113	109/116		5.1 %	0.49 [0.19, 1.29]
Hnsgen 1996	1/53	2/55		0.9 %	0.51 [0.04, 5.79]
Kalb 2001	3/37	2/35		1.5 %	1.46 [0.23, 9.28]
Kasper 2006	99/250	22/82		13.1 %	1.79 [1.03, 3.10]
Laakmann 1998	14/49	15/49		6.1 %	0.91 [0.38, 2.16]



(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	hypericum n/N	placebo n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Lecrubier 2002	57/186	70/189		18.7 %	0.75 [0.49, 1.15]
Philipp 1999	23/106	9/47		6.2 %	1.17 [0.49, 2.77]
Schrader 1998	6/80	5/79		3.2 %	1.20 [0.35, 4.10]
Uebelhack 2004	14/70	7/70		4.9 %	2.25 [0.85, 5.97]
Volz 2000	12/70	19/70		6.8 %	0.56 [0.25, 1.25]
Witte 1995	0/48	1/49		0.5 %	0.33 [0.01, 8.39]
Total (95% CI)	**1359**	**1137**		**100.0 %**	**0.98 [0.78, 1.23]**
Total events: 418 (hypericum), 353 (placebo)					
Heterogeneity: Tau^2 = 0.02; Chi^2 = 14.98, df = 13 (P = 0.31); I^2 =13%					
Test for overall effect: Z = 0.21 (P = 0.84)					

Analysis 4.1. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 1 Responder (intent to treat) - primary analysis.

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

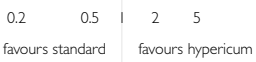
Outcome: 1 Responder (intent to treat) - primary analysis

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 vs. older antidepressants					
Harrer 1993	27/51	28/51		4.3 %	0.96 [0.67, 1.38]
Philipp 1999	76/106	70/110		12.2 %	1.13 [0.94, 1.36]
Vorbach 1997	36/107	41/102		4.4 %	0.84 [0.59, 1.20]
Wheatley 1997	40/87	42/78		5.7 %	0.85 [0.63, 1.16]
Woelk 2000	68/157	67/167		7.6 %	1.08 [0.83, 1.40]
Subtotal (95% CI)	**508**	**508**		**34.2 %**	**1.02 [0.90, 1.15]**
Total events: 247 (hypericum), 248 (standard)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 3.97, df = 4 (P = 0.41); I^2 =0.0%					
Test for overall effect: Z = 0.30 (P = 0.77)					
2 vs. SSRIs					
Behnke 2002	16/35	21/35		2.9 %	0.76 [0.49, 1.20]

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Bjerkenstedt 2005	22/54	20/54		2.6 %	1.10 [0.68, 1.77]
Brenner 2000	7/15	6/15		0.9 %	1.17 [0.51, 2.66]
Fava 2005	17/45	14/47		1.8 %	1.27 [0.71, 2.26]
Gastpar 2005	70/123	72/118		10.2 %	0.93 [0.76, 1.15]
Gastpar 2006	71/131	71/127		9.6 %	0.97 [0.78, 1.21]
Harrer 1999	50/77	57/84		9.6 %	0.96 [0.77, 1.19]
HDTSG 2002	46/113	55/109		6.2 %	0.81 [0.60, 1.08]
Moreno 2005	4/20	11/20		0.7 %	0.36 [0.14, 0.95]
Schrader 2000	57/125	39/113		5.3 %	1.32 [0.96, 1.82]
Szegedi 2005	86/122	73/122		12.2 %	1.18 [0.98, 1.42]
van Gurp 2002	25/45	23/45		3.8 %	1.09 [0.74, 1.60]
Subtotal (95% CI)	**905**	**889**		**65.8 %**	**1.00 [0.90, 1.12]**
Total events: 471 (hypericum), 462 (standard)					
Heterogeneity: Tau^2 = 0.01; Chi^2 = 15.43, df = 11 (P = 0.16); I^2 =29%					
Test for overall effect: Z = 0.06 (P = 0.95)					
Total (95% CI)	**1413**	**1397**		**100.0 %**	**1.01 [0.93, 1.09]**
Total events: 718 (hypericum), 710 (standard)					
Heterogeneity: Tau^2 = 0.00; Chi^2 = 19.37, df = 16 (P = 0.25); I^2 =17%					
Test for overall effect: Z = 0.14 (P = 0.89)					



Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 1 Responder (intent to treat) - primary analysis

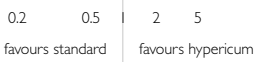
Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
1 vs. older antidepressants				
Harrer 1993	27/51	28/51		0.96 [0.67, 1.38]
Philipp 1999	76/106	70/110		1.13 [0.94, 1.36]
Vorbach 1997	36/107	41/102		0.84 [0.59, 1.20]
Wheatley 1997	40/87	42/78		0.85 [0.63, 1.16]
Woelk 2000	68/157	67/167		1.08 [0.83, 1.40]
Subtotal (95% CI)	**508**	**508**		**1.02 [0.90, 1.15]**
Total events: 247 (hypericum), 248 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.97$, df = 4 (P = 0.41); $I^2 = 0.0\%$				
Test for overall effect: Z = 0.30 (P = 0.77)				

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 1 Responder (intent to treat) - primary analysis

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 vs. SSRIs				
Behnke 2002	16/35	21/35		0.76 [0.49, 1.20]
Bjerkenstedt 2005	22/54	20/54		1.10 [0.68, 1.77]
Brenner 2000	7/15	6/15		1.17 [0.51, 2.66]
Fava 2005	17/45	14/47		1.27 [0.71, 2.26]
Gastpar 2005	70/123	72/118		0.93 [0.76, 1.15]
Gastpar 2006	71/131	71/127		0.97 [0.78, 1.21]
Harrer 1999	50/77	57/84		0.96 [0.77, 1.19]
HDTSG 2002	46/113	55/109		0.81 [0.60, 1.08]
Moreno 2005	4/20	11/20		0.36 [0.14, 0.95]



(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
Schrader 2000	57/125	39/113		1.32 [0.96, 1.82]
Szegedi 2005	86/122	73/122		1.18 [0.98, 1.42]
van Gurp 2002	25/45	23/45		1.09 [0.74, 1.60]
Subtotal (95% CI)	**905**	**889**		**1.00 [0.90, 1.12]**
Total events: 471 (hypericum), 462 (standard)				
Heterogeneity: Tau^2 = 0.01; Chi^2 = 15.43, df = 11 (P = 0.16); I^2 =29%				
Test for overall effect: Z = 0.06 (P = 0.95)				

Analysis 4.2. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 2 Responder (per protocol).

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

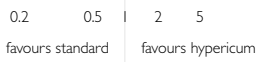
Outcome: 2 Responder (per protocol)

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 vs. older antidepressants					
Harrer 1993	27/44	28/42		5.0 %	0.92 [0.67, 1.26]
Philipp 1999	76/93	70/99		10.2 %	1.16 [0.99, 1.36]
Vorbach 1997	36/98	41/88		4.5 %	0.79 [0.56, 1.11]
Wheatley 1997	40/67	42/54		7.0 %	0.77 [0.60, 0.98]
Woelk 2000	68/138	67/131		7.2 %	0.96 [0.76, 1.22]
Subtotal (95% CI)	**440**	**414**		**33.9 %**	**0.93 [0.78, 1.11]**
Total events: 247 (hypericum), 248 (standard)					
Heterogeneity: Tau^2 = 0.02; Chi^2 = 10.10, df = 4 (P = 0.04); I^2 =60%					
Test for overall effect: Z = 0.81 (P = 0.42)					
2 vs. SSRIs					
Behnke 2002	16/29	21/32		3.4 %	0.84 [0.56, 1.27]
Bjerkenstedt 2005	22/48	20/48		2.9 %	1.10 [0.70, 1.73]
Brenner 2000	7/8	6/12		1.7 %	1.75 [0.94, 3.26]
Fava 2005	17/27	14/21		3.4 %	0.94 [0.62, 1.44]

(*Continued* . . .)

(. . . *Continued*)

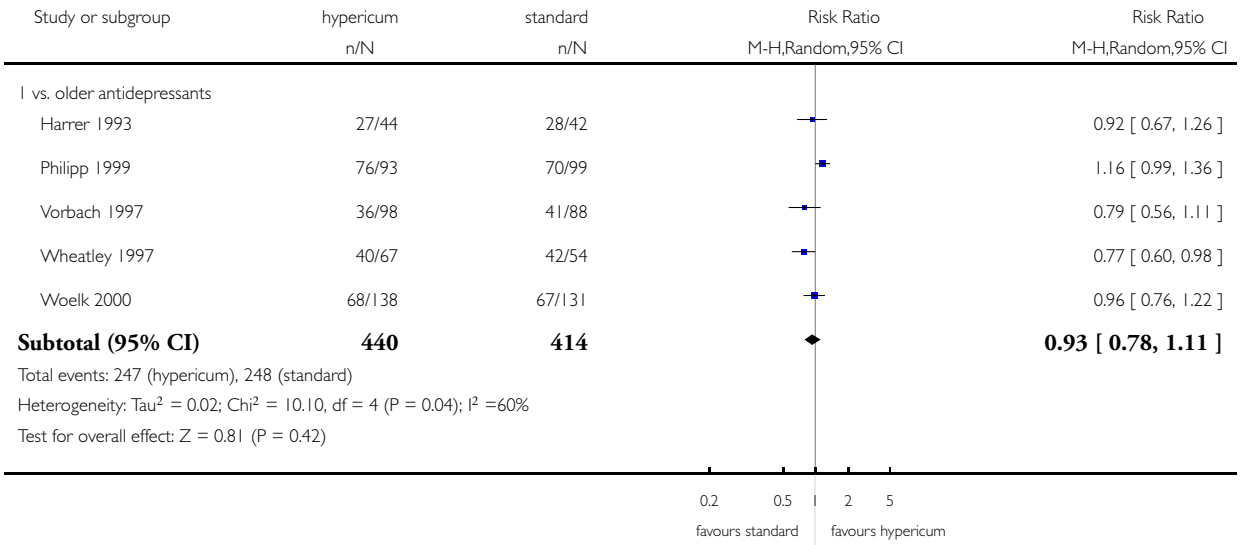
Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Gastpar 2005	70/102	72/98		9.4 %	0.93 [0.78, 1.12]
Gastpar 2006	71/103	71/104		9.1 %	1.01 [0.84, 1.21]
Harrer 1999	50/69	57/68		9.3 %	0.86 [0.72, 1.03]
HDTSG 2002	46/82	55/84		6.9 %	0.86 [0.67, 1.10]
Moreno 2005	4/18	11/16		0.8 %	0.32 [0.13, 0.82]
Schrader 2000	57/125	39/113		5.0 %	1.32 [0.96, 1.82]
Szegedi 2005	70/97	60/91		8.8 %	1.09 [0.90, 1.33]
van Gurp 2002	20/29	22/28		5.2 %	0.88 [0.64, 1.20]
Subtotal (95% CI)	**737**	**715**		**66.1 %**	**0.97 [0.87, 1.08]**
Total events: 450 (hypericum), 448 (standard)					
Heterogeneity: Tau^2 = 0.01; Chi^2 = 18.17, df = 11 (P = 0.08); I^2 =39%					
Test for overall effect: Z = 0.55 (P = 0.59)					
Total (95% CI)	**1177**	**1129**		**100.0 %**	**0.96 [0.88, 1.05]**
Total events: 697 (hypericum), 696 (standard)					
Heterogeneity: Tau^2 = 0.01; Chi^2 = 27.83, df = 16 (P = 0.03); I^2 =43%					
Test for overall effect: Z = 0.92 (P = 0.36)					



Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

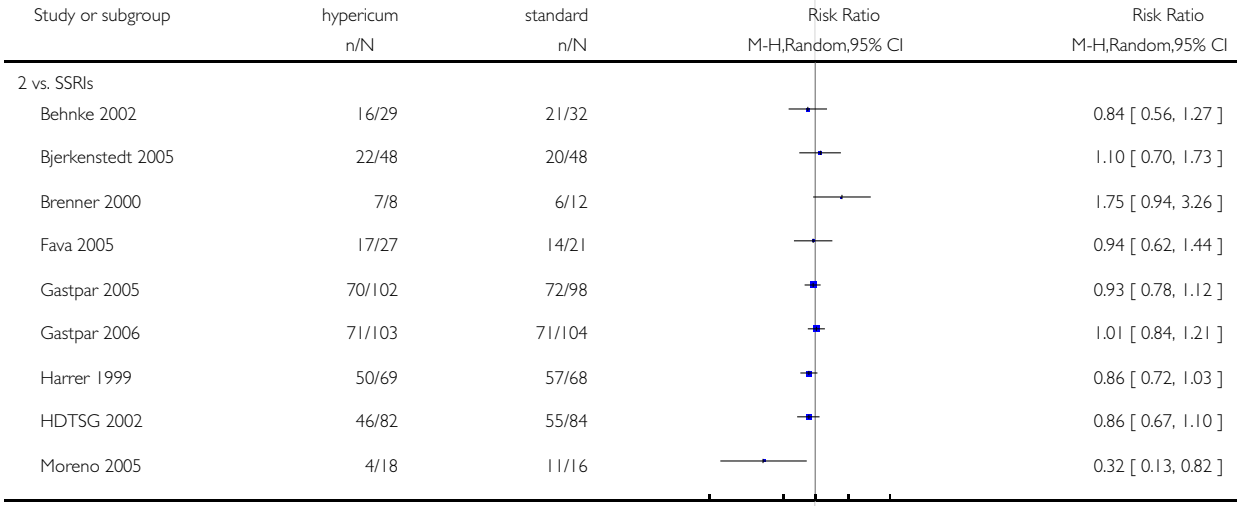
Outcome: 2 Responder (per protocol)



Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 2 Responder (per protocol)



(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
Schrader 2000	57/125	39/113		1.32 [0.96, 1.82]
Szegedi 2005	70/97	60/91		1.09 [0.90, 1.33]
van Gurp 2002	20/29	22/28		0.88 [0.64, 1.20]
Subtotal (95% CI)	**737**	**715**		**0.97 [0.87, 1.08]**
Total events: 450 (hypericum), 448 (standard)				
Heterogeneity: Tau^2 = 0.01; Chi^2 = 18.17, df = 11 (P = 0.08); I^2 =39%				
Test for overall effect: Z = 0.55 (P = 0.59)				

0.2 0.5 1 2 5
favours standard favours hypericum

Analysis 4.3. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 3 Responders according to CGI (Clinical Global Impression Index at least "much improved").

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 3 Responders according to CGI (Clinical Global Impression Index at least "much improved")

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 vs. older antidepressants					
Harrer 1993	22/51	20/51		2.5 %	1.10 [0.69, 1.75]
Philipp 1999	77/106	78/110		13.2 %	1.02 [0.87, 1.21]
Vorbach 1997	60/107	64/102		8.8 %	0.89 [0.71, 1.12]
Wheatley 1997	55/87	53/78		9.0 %	0.93 [0.75, 1.16]
Subtotal (95% CI)	**351**	**341**		**33.6 %**	**0.97 [0.87, 1.09]**
Total events: 214 (hypericum), 215 (standard)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 1.34, df = 3 (P = 0.72); I^2 =0.0%					
Test for overall effect: Z = 0.52 (P = 0.60)					
2 vs. newer antidepressants					
Behnke 2002	21/35	24/35		4.2 %	0.88 [0.62, 1.24]
Bjerkenstedt 2005	25/54	22/54		2.9 %	1.14 [0.74, 1.75]
Gastpar 2005	70/123	62/118		8.5 %	1.08 [0.86, 1.36]
Gastpar 2006	85/131	86/127		12.6 %	0.96 [0.81, 1.14]

0.2 0.5 1 2 5
favours standard favours hypericum

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Harrer 1999	49/77	54/84		8.4 %	0.99 [0.78, 1.25]
HDTSG 2002	47/113	61/109		6.4 %	0.74 [0.56, 0.98]
Schrader 2000	94/125	71/113		12.6 %	1.20 [1.01, 1.42]
Szegedi 2005	83/122	70/122		10.8 %	1.19 [0.98, 1.44]
Subtotal (95% CI)	**780**	**762**		**66.4 %**	**1.03 [0.92, 1.15]**
Total events: 474 (hypericum), 450 (standard)					
Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 12.28$, df = 7 (P = 0.09); $I^2 = 43\%$					
Test for overall effect: Z = 0.45 (P = 0.66)					
Total (95% CI)	**1131**	**1103**		**100.0 %**	**1.01 [0.94, 1.09]**
Total events: 688 (hypericum), 665 (standard)					
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 14.54$, df = 11 (P = 0.20); $I^2 = 24\%$					
Test for overall effect: Z = 0.25 (P = 0.80)					

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 3 Responders according to CGI (Clinical Global Impression Index at least "much improved")

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
1 vs. older antidepressants				
Harrer 1993	22/51	20/51		1.10 [0.69, 1.75]
Philipp 1999	77/106	78/110		1.02 [0.87, 1.21]
Vorbach 1997	60/107	64/102		0.89 [0.71, 1.12]
Wheatley 1997	55/87	53/78		0.93 [0.75, 1.16]
Subtotal (95% CI)	**351**	**341**		**0.97 [0.87, 1.09]**
Total events: 214 (hypericum), 215 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 1.34$, df = 3 (P = 0.72); $I^2 = 0.0\%$				
Test for overall effect: Z = 0.52 (P = 0.60)				

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 3 Responders according to CGI (Clinical Global Impression Index at least "much improved")

Study or subgroup	hypericum n/N	standard n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 vs. newer antidepressants				
Behnke 2002	21/35	24/35		0.88 [0.62, 1.24]
Bjerkenstedt 2005	25/54	22/54		1.14 [0.74, 1.75]
Gastpar 2005	70/123	62/118		1.08 [0.86, 1.36]
Gastpar 2006	85/131	86/127		0.96 [0.81, 1.14]
Harrer 1999	49/77	54/84		0.99 [0.78, 1.25]
HDTSG 2002	47/113	61/109		0.74 [0.56, 0.98]
Schrader 2000	94/125	71/113		1.20 [1.01, 1.42]
Szegedi 2005	83/122	70/122		1.19 [0.98, 1.44]
Subtotal (95% CI)	**780**	**762**		**1.03 [0.92, 1.15]**
Total events: 474 (hypericum), 450 (standard)				
Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 12.28$, df = 7 (P = 0.09); $I^2 = 43\%$				
Test for overall effect: Z = 0.45 (P = 0.66)				

Analysis 4.4. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 4 Responder among studies from German-speaking studies and other studies.

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 4 Responder among studies from German-speaking studies and other studies

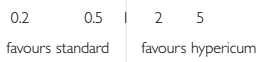
Study or subgroup	hypericum n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 Studies from German-speaking countries					
Gastpar 2005	70/102	72/98		12.6 %	0.93 [0.78, 1.12]
Gastpar 2006	71/131	71/127		9.3 %	0.97 [0.78, 1.21]
Harrer 1993	27/51	28/51		4.2 %	0.96 [0.67, 1.38]
Harrer 1999	50/77	57/84		9.3 %	0.96 [0.77, 1.19]
Philipp 1999	76/106	70/110		11.9 %	1.13 [0.94, 1.36]
Schrader 2000	57/125	39/113		5.2 %	1.32 [0.96, 1.82]
Szegedi 2005	86/122	73/122		11.9 %	1.18 [0.98, 1.42]
Vorbach 1997	36/107	41/102		4.3 %	0.84 [0.59, 1.20]
Woelk 2000	68/157	67/167		7.4 %	1.08 [0.83, 1.40]
Subtotal (95% CI)	**978**	**974**		**76.0 %**	**1.04 [0.96, 1.13]**
Total events: 541 (hypericum), 518 (Control)					
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 8.67$, df = 8 (P = 0.37); $I^2 = 8\%$					
Test for overall effect: Z = 0.95 (P = 0.34)					
2 Studies from other countries					
Behnke 2002	16/35	21/35		2.8 %	0.76 [0.49, 1.20]
Bjerkenstedt 2005	22/54	20/54		2.6 %	1.10 [0.68, 1.77]
Brenner 2000	7/15	6/15		0.9 %	1.17 [0.51, 2.66]
Fava 2005	17/45	14/47		1.8 %	1.27 [0.71, 2.26]
HDTSG 2002	46/113	55/109		6.1 %	0.81 [0.60, 1.08]
Moreno 2005	4/20	11/20		0.7 %	0.36 [0.14, 0.95]
van Gurp 2002	25/45	23/45		3.7 %	1.09 [0.74, 1.60]
Wheatley 1997	40/87	42/78		5.5 %	0.85 [0.63, 1.16]
Subtotal (95% CI)	**414**	**403**		**24.0 %**	**0.90 [0.76, 1.06]**
Total events: 177 (hypericum), 192 (Control)					
Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 7.93$, df = 7 (P = 0.34); $I^2 = 12\%$					
Test for overall effect: Z = 1.22 (P = 0.22)					
Total (95% CI)	**1392**	**1377**		**100.0 %**	**1.00 [0.93, 1.09]**
Total events: 718 (hypericum), 710 (Control)					
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 19.53$, df = 16 (P = 0.24); $I^2 = 18\%$					
Test for overall effect: Z = 0.10 (P = 0.92)					

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 4 Responder among studies from German-speaking studies and other studies

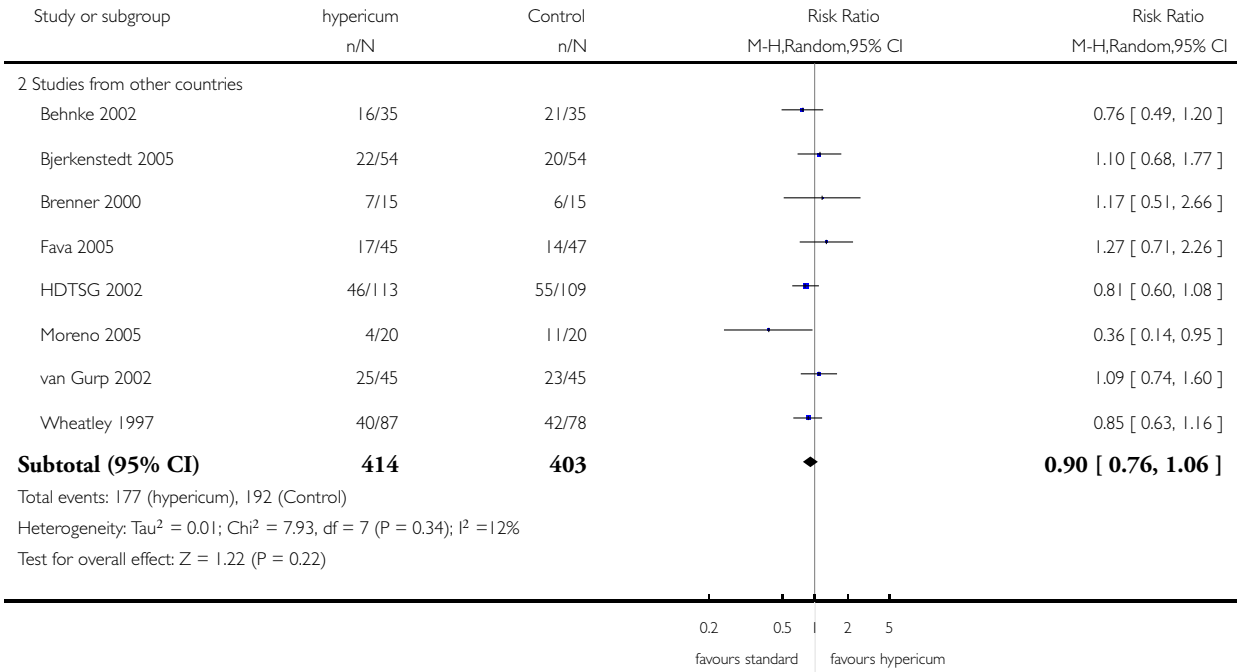
Study or subgroup	hypericum n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
1 Studies from German-speaking countries				
Gastpar 2005	70/102	72/98		0.93 [0.78, 1.12]
Gastpar 2006	71/131	71/127		0.97 [0.78, 1.21]
Harrer 1993	27/51	28/51		0.96 [0.67, 1.38]
Harrer 1999	50/77	57/84		0.96 [0.77, 1.19]
Philipp 1999	76/106	70/110		1.13 [0.94, 1.36]
Schrader 2000	57/125	39/113		1.32 [0.96, 1.82]
Szegedi 2005	86/122	73/122		1.18 [0.98, 1.42]
Vorbach 1997	36/107	41/102		0.84 [0.59, 1.20]
Woelk 2000	68/157	67/167		1.08 [0.83, 1.40]
Subtotal (95% CI)	**978**	**974**		**1.04 [0.96, 1.13]**
Total events: 541 (hypericum), 518 (Control)				
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 8.67$, df = 8 (P = 0.37); $I^2 = 8\%$				
Test for overall effect: Z = 0.95 (P = 0.34)				



Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 4 Responder among studies from German-speaking studies and other studies



Analysis 4.5. Comparison 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures, Outcome 5 Remission (HAMD score < 8).

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 5 Remission (HAMD score < 8)

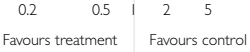
Study or subgroup	Treatment n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
1 vs. older antidepressants					
Subtotal (95% CI)	**0**	**0**		**0.0 %**	**Not estimable**
Total events: 0 (Treatment), 0 (Control)					
Heterogeneity: not applicable					
Test for overall effect: not applicable					
2 vs. SSRIs					
Bjerkenstedt 2005	13/54	15/54		9.0 %	0.87 [0.46, 1.64]
Fava 2005	17/45	14/47		11.1 %	1.27 [0.71, 2.26]
Gastpar 2005	53/123	44/118		38.5 %	1.16 [0.85, 1.58]
Szegedi 2005	61/122	43/122		41.4 %	1.42 [1.05, 1.91]
Subtotal (95% CI)	**344**	**341**		**100.0 %**	**1.24 [1.02, 1.50]**
Total events: 144 (Treatment), 116 (Control)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 2.19, df = 3 (P = 0.53); I^2 =0.0%					
Test for overall effect: Z = 2.18 (P = 0.029)					
Total (95% CI)	**344**	**341**		**100.0 %**	**1.24 [1.02, 1.50]**
Total events: 144 (Treatment), 116 (Control)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 2.19, df = 3 (P = 0.53); I^2 =0.0%					
Test for overall effect: Z = 2.18 (P = 0.029)					

Review: St John's wort for major depression

Comparison: 4 Hypericum mono-preparations vs. standard antidepressants. A. Dichotomous measures

Outcome: 5 Remission (HAMD score < 8)

Study or subgroup	Treatment n/N	Control n/N	Risk Ratio M-H,Random,95% CI	Risk Ratio M-H,Random,95% CI
2 vs. SSRIs				
Bjerkenstedt 2005	13/54	15/54		0.87 [0.46, 1.64]
Fava 2005	17/45	14/47		1.27 [0.71, 2.26]
Gastpar 2005	53/123	44/118		1.16 [0.85, 1.58]
Szegedi 2005	61/122	43/122		1.42 [1.05, 1.91]
Subtotal (95% CI)	**344**	**341**		**1.24 [1.02, 1.50]**
Total events: 144 (Treatment), 116 (Control)				
Heterogeneity: Tau^2 = 0.0; Chi^2 = 2.19, df = 3 (P = 0.53); I^2 =0.0%				
Test for overall effect: Z = 2.18 (P = 0.029)				



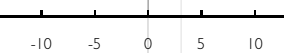
Analysis 5.1. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 1 Mean HAMD (Hamilton Rating Scale for Depression) after therapy.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 1 Mean HAMD (Hamilton Rating Scale for Depression) after therapy

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Harrer 1993	44	11.2 (9.9)	42	10.5 (7.5)		4.1 %	0.70 [-3.00, 4.40]
Philipp 1999	100	9.2 (6.2)	105	10.6 (6.1)		11.0 %	-1.40 [-3.08, 0.28]
Vorbach 1997	98	14.4 (6.1)	88	13.4 (5.9)		10.7 %	1.00 [-0.73, 2.73]
Subtotal (95% CI)	**242**		**235**			**25.8 %**	**-0.06 [-1.82, 1.71]**
Heterogeneity: Tau^2 = 1.19; Chi^2 = 4.02, df = 2 (P = 0.13); I^2 =50%							
Test for overall effect: Z = 0.06 (P = 0.95)							
2 vs. SSRIs							
Bjerkenstedt 2005	54	15 (8.4)	54	14.9 (8.4)		5.2 %	0.10 [-3.07, 3.27]
Brenner 2000	13	12.7 (6.7)	15	12.5 (5.6)		2.8 %	0.20 [-4.41, 4.81]
Fava 2005	45	10.2 (6.6)	47	13.3 (7.3)		6.0 %	-3.10 [-5.94, -0.26]



(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Gastpar 2005	101	8.3 (5.5)	97	8.1 (5.6)		11.8 %	0.20 [-1.35, 1.75]
Gastpar 2006	131	10.3 (6.4)	127	10.3 (6.4)		11.7 %	0.0 [-1.56, 1.56]
HDTSG 2002	82	12.9 (7.1)	77	10.7 (5.9)		9.2 %	2.20 [0.18, 4.22]
Schrader 2000	125	10.92 (4.27)	113	11.47 (4.04)		15.0 %	-0.55 [-1.61, 0.51]
Szegedi 2005	122	11.2 (9)	122	14.2 (8.9)		8.2 %	-3.00 [-5.25, -0.75]
van Gurp 2002	44	9.43 (8.26)	43	11.56 (8.41)		4.4 %	-2.13 [-5.63, 1.37]
Subtotal (95% CI)	**717**		**695**			**74.2 %**	**-0.52 [-1.55, 0.51]**
Heterogeneity: Tau² = 1.16; Chi² = 16.96, df = 8 (P = 0.03); I² =53%							
Test for overall effect: Z = 0.99 (P = 0.32)							
Total (95% CI)	**959**		**930**			**100.0 %**	**-0.39 [-1.23, 0.45]**
Heterogeneity: Tau² = 0.94; Chi² = 21.13, df = 11 (P = 0.03); I² =48%							
Test for overall effect: Z = 0.91 (P = 0.36)							

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 1 Mean HAMD (Hamilton Rating Scale for Depression) after therapy

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 vs. older antidepressants						
Harrer 1993	44	11.2 (9.9)	42	10.5 (7.5)		0.70 [-3.00, 4.40]
Philipp 1999	100	9.2 (6.2)	105	10.6 (6.1)		-1.40 [-3.08, 0.28]
Vorbach 1997	98	14.4 (6.1)	88	13.4 (5.9)		1.00 [-0.73, 2.73]
Subtotal (95% CI)	**242**		**235**			**-0.06 [-1.82, 1.71]**
Heterogeneity: Tau² = 1.19; Chi² = 4.02, df = 2 (P = 0.13); I² =50%						
Test for overall effect: Z = 0.06 (P = 0.95)						

Review: St John's wort for major depression

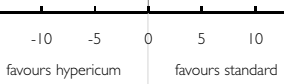
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 1 Mean HAMD (Hamilton Rating Scale for Depression) after therapy

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Bjerkenstedt 2005	54	15 (8.4)	54	14.9 (8.4)		0.10 [-3.07, 3.27]
Brenner 2000	13	12.7 (6.7)	15	12.5 (5.6)		0.20 [-4.41, 4.81]
Fava 2005	45	10.2 (6.6)	47	13.3 (7.3)		-3.10 [-5.94, -0.26]
Gastpar 2005	101	8.3 (5.5)	97	8.1 (5.6)		0.20 [-1.35, 1.75]
Gastpar 2006	131	10.3 (6.4)	127	10.3 (6.4)		0.0 [-1.56, 1.56]
HDTSG 2002	82	12.9 (7.1)	77	10.7 (5.9)		2.20 [0.18, 4.22]
Schrader 2000	125	10.92 (4.27)	113	11.47 (4.04)		-0.55 [-1.61, 0.51]
Szegedi 2005	122	11.2 (9)	122	14.2 (8.9)		-3.00 [-5.25, -0.75]
van Gurp 2002	44	9.43 (8.26)	43	11.56 (8.41)		-2.13 [-5.63, 1.37]
Subtotal (95% CI)	**717**		**695**			**-0.52 [-1.55, 0.51]**

Heterogeneity: $Tau^2 = 1.16$; $Chi^2 = 16.96$, df = 8 (P = 0.03); $I^2 = 53\%$

Test for overall effect: Z = 0.99 (P = 0.32)



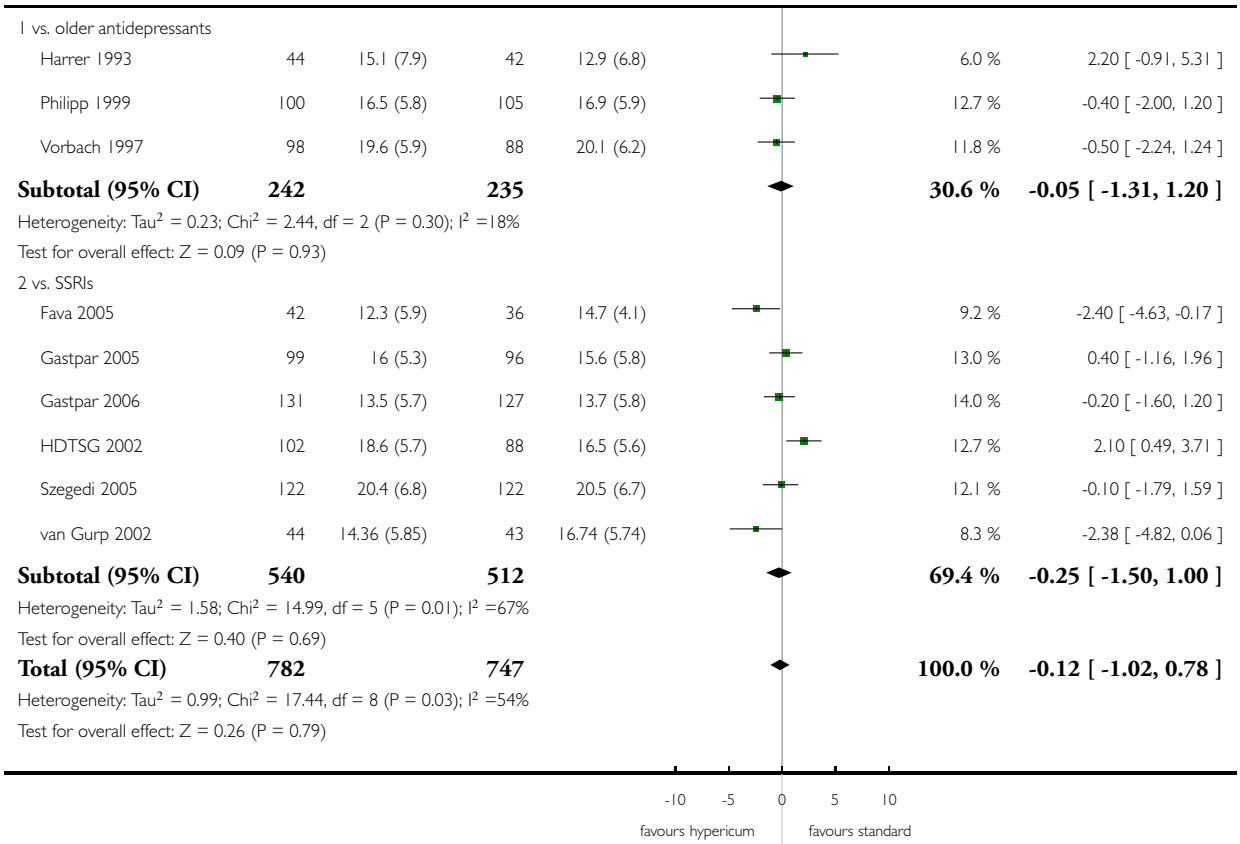
Analysis 5.2. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 or 3 weeks of treatment.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 or 3 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Harrer 1993	44	15.1 (7.9)	42	12.9 (6.8)		6.0 %	2.20 [-0.91, 5.31]
Philipp 1999	100	16.5 (5.8)	105	16.9 (5.9)		12.7 %	-0.40 [-2.00, 1.20]
Vorbach 1997	98	19.6 (5.9)	88	20.1 (6.2)		11.8 %	-0.50 [-2.24, 1.24]
Subtotal (95% CI)	**242**		**235**			**30.6 %**	**-0.05 [-1.31, 1.20]**
Heterogeneity: $Tau^2 = 0.23$; $Chi^2 = 2.44$, df = 2 (P = 0.30); $I^2 = 18\%$							
Test for overall effect: Z = 0.09 (P = 0.93)							
2 vs. SSRIs							
Fava 2005	42	12.3 (5.9)	36	14.7 (4.1)		9.2 %	-2.40 [-4.63, -0.17]
Gastpar 2005	99	16 (5.3)	96	15.6 (5.8)		13.0 %	0.40 [-1.16, 1.96]
Gastpar 2006	131	13.5 (5.7)	127	13.7 (5.8)		14.0 %	-0.20 [-1.60, 1.20]
HDTSG 2002	102	18.6 (5.7)	88	16.5 (5.6)		12.7 %	2.10 [0.49, 3.71]
Szegedi 2005	122	20.4 (6.8)	122	20.5 (6.7)		12.1 %	-0.10 [-1.79, 1.59]
van Gurp 2002	44	14.36 (5.85)	43	16.74 (5.74)		8.3 %	-2.38 [-4.82, 0.06]
Subtotal (95% CI)	**540**		**512**			**69.4 %**	**-0.25 [-1.50, 1.00]**
Heterogeneity: $Tau^2 = 1.58$; $Chi^2 = 14.99$, df = 5 (P = 0.01); $I^2 = 67\%$							
Test for overall effect: Z = 0.40 (P = 0.69)							
Total (95% CI)	**782**		**747**			**100.0 %**	**-0.12 [-1.02, 0.78]**
Heterogeneity: $Tau^2 = 0.99$; $Chi^2 = 17.44$, df = 8 (P = 0.03); $I^2 = 54\%$							
Test for overall effect: Z = 0.26 (P = 0.79)							



Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 or 3 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 vs. older antidepressants						
Harrer 1993	44	15.1 (7.9)	42	12.9 (6.8)		2.20 [-0.91, 5.31]
Philipp 1999	100	16.5 (5.8)	105	16.9 (5.9)		-0.40 [-2.00, 1.20]
Vorbach 1997	98	19.6 (5.9)	88	20.1 (6.2)		-0.50 [-2.24, 1.24]
Subtotal (95% CI)	**242**		**235**			**-0.05 [-1.31, 1.20]**

Heterogeneity: Tau² = 0.23; Chi² = 2.44, df = 2 (P = 0.30); I² =18%

Test for overall effect: Z = 0.09 (P = 0.93)

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 2 Mean HAMD (Hamilton Rating Scale for Depression) scores after 2 or 3 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Fava 2005	42	12.3 (5.9)	36	14.7 (4.1)		-2.40 [-4.63, -0.17]
Gastpar 2005	99	16 (5.3)	96	15.6 (5.8)		0.40 [-1.16, 1.96]
Gastpar 2006	131	13.5 (5.7)	127	13.7 (5.8)		-0.20 [-1.60, 1.20]
HDTSG 2002	102	18.6 (5.7)	88	16.5 (5.6)		2.10 [0.49, 3.71]
Szegedi 2005	122	20.4 (6.8)	122	20.5 (6.7)		-0.10 [-1.79, 1.59]
van Gurp 2002	44	14.36 (5.85)	43	16.74 (5.74)		-2.38 [-4.82, 0.06]
Subtotal (95% CI)	**540**		**512**			**-0.25 [-1.50, 1.00]**

Heterogeneity: Tau² = 1.58; Chi² = 14.99, df = 5 (P = 0.01); I² =67%

Test for overall effect: Z = 0.40 (P = 0.69)

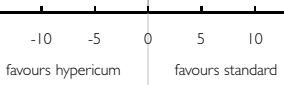
Analysis 5.3. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 3 Mean HAMD (Hamilton Rating Scale for Depression) scores after 4 weeks of treatment.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 3 Mean HAMD (Hamilton Rating Scale for Depression) scores after 4 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Harrer 1993	44	11.2 (9.9)	42	10.5 (7.5)		6.5 %	0.70 [-3.00, 4.40]
Philipp 1999	100	12.3 (5.9)	105	13 (5.6)		14.8 %	-0.70 [-2.28, 0.88]
Vorbach 1997	98	17.1 (6.6)	88	16.3 (5.9)		13.6 %	0.80 [-1.00, 2.60]
Subtotal (95% CI)	**242**		**235**			**34.9 %**	**0.02 [-1.11, 1.15]**
Heterogeneity: Tau² = 0.0; Chi² = 1.66, df = 2 (P = 0.44); I² =0.0%							
Test for overall effect: Z = 0.04 (P = 0.97)							
2 vs. SSRIs							
Bjerkenstedt 2005	54	15 (8.4)	54	14.9 (8.4)		7.9 %	0.10 [-3.07, 3.27]
Fava 2005	35	11.5 (6.9)	35	14.3 (4.7)		9.3 %	-2.80 [-5.57, -0.03]
Gastpar 2005	100	12.1 (5.7)	97	12 (6)		14.4 %	0.10 [-1.54, 1.74]
HDTSG 2002	97	16.9 (7.1)	87	14.5 (6.1)		13.0 %	2.40 [0.49, 4.31]
Szegedi 2005	122	15.1 (8.3)	122	16.3 (8.4)		12.1 %	-1.20 [-3.30, 0.90]
van Gurp 2002	44	10.93 (7.03)	43	14.79 (7.55)		8.3 %	-3.86 [-6.93, -0.79]
Subtotal (95% CI)	**452**		**438**			**65.1 %**	**-0.69 [-2.44, 1.06]**
Heterogeneity: Tau² = 3.25; Chi² = 17.01, df = 5 (P = 0.004); I² =71%							
Test for overall effect: Z = 0.78 (P = 0.44)							
Total (95% CI)	**694**		**673**			**100.0 %**	**-0.34 [-1.48, 0.80]**
Heterogeneity: Tau² = 1.65; Chi² = 18.83, df = 8 (P = 0.02); I² =58%							
Test for overall effect: Z = 0.58 (P = 0.56)							



Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 3 Mean HAMD (Hamilton Rating Scale for Depression) scores after 4 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 vs. older antidepressants						
Harrer 1993	44	11.2 (9.9)	42	10.5 (7.5)		0.70 [-3.00, 4.40]
Philipp 1999	100	12.3 (5.9)	105	13 (5.6)		-0.70 [-2.28, 0.88]
Vorbach 1997	98	17.1 (6.6)	88	16.3 (5.9)		0.80 [-1.00, 2.60]
Subtotal (95% CI)	**242**		**235**			**0.02 [-1.11, 1.15]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 1.66$, df = 2 (P = 0.44); $I^2 = 0.0\%$

Test for overall effect: Z = 0.04 (P = 0.97)

Review: St John's wort for major depression

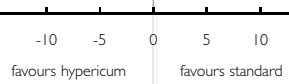
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 3 Mean HAMD (Hamilton Rating Scale for Depression) scores after 4 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Bjerkenstedt 2005	54	15 (8.4)	54	14.9 (8.4)		0.10 [-3.07, 3.27]
Fava 2005	35	11.5 (6.9)	35	14.3 (4.7)		-2.80 [-5.57, -0.03]
Gastpar 2005	100	12.1 (5.7)	97	12 (6)		0.10 [-1.54, 1.74]
HDTSG 2002	97	16.9 (7.1)	87	14.5 (6.1)		2.40 [0.49, 4.31]
Szegedi 2005	122	15.1 (8.3)	122	16.3 (8.4)		-1.20 [-3.30, 0.90]
van Gurp 2002	44	10.93 (7.03)	43	14.79 (7.55)		-3.86 [-6.93, -0.79]
Subtotal (95% CI)	**452**		**438**			**-0.69 [-2.44, 1.06]**

Heterogeneity: $Tau^2 = 3.25$; $Chi^2 = 17.01$, df = 5 (P = 0.004); $I^2 = 71\%$

Test for overall effect: Z = 0.78 (P = 0.44)



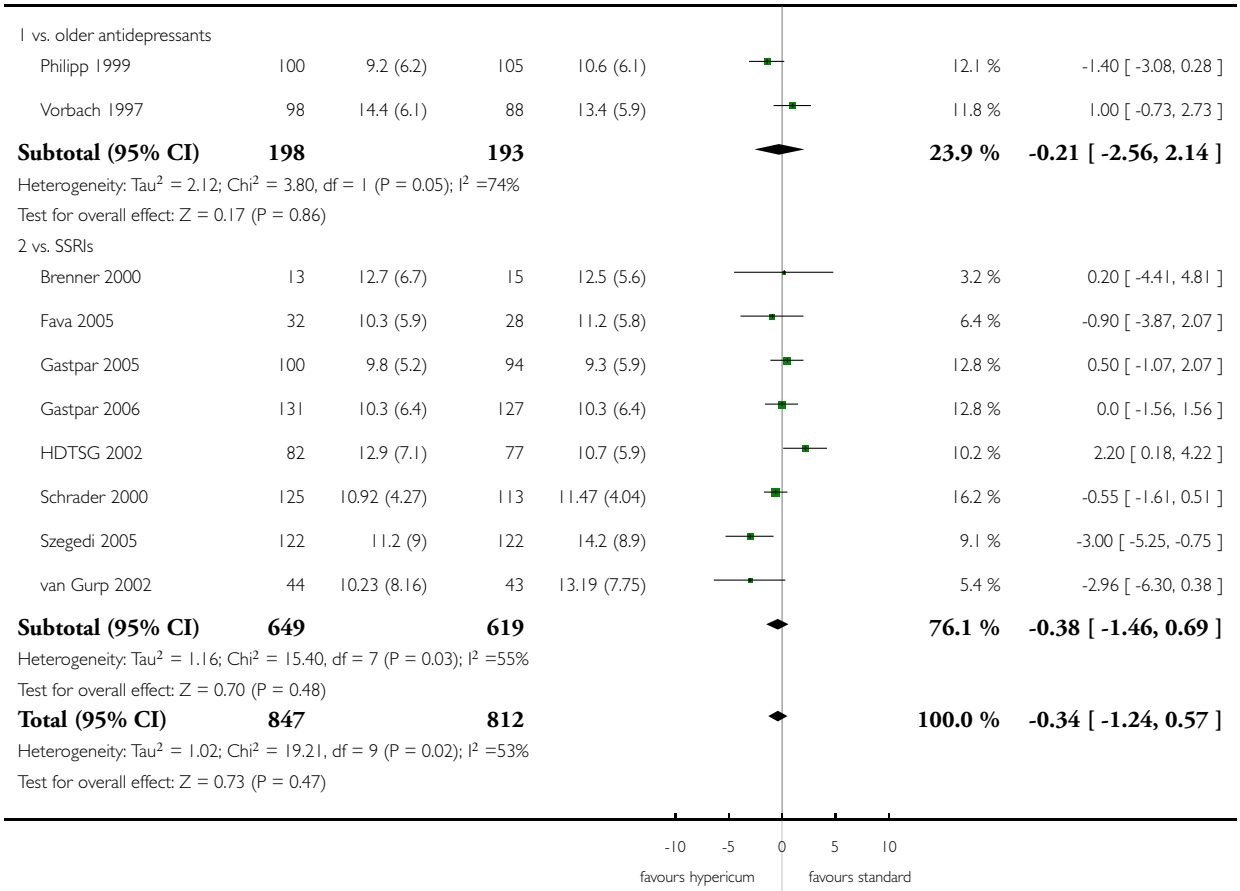
Analysis 5.4. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Philipp 1999	100	9.2 (6.2)	105	10.6 (6.1)		12.1 %	-1.40 [-3.08, 0.28]
Vorbach 1997	98	14.4 (6.1)	88	13.4 (5.9)		11.8 %	1.00 [-0.73, 2.73]
Subtotal (95% CI)	**198**		**193**			**23.9 %**	**-0.21 [-2.56, 2.14]**
Heterogeneity: $Tau^2 = 2.12$; $Chi^2 = 3.80$, df = 1 (P = 0.05); $I^2 = 74\%$							
Test for overall effect: Z = 0.17 (P = 0.86)							
2 vs. SSRIs							
Brenner 2000	13	12.7 (6.7)	15	12.5 (5.6)		3.2 %	0.20 [-4.41, 4.81]
Fava 2005	32	10.3 (5.9)	28	11.2 (5.8)		6.4 %	-0.90 [-3.87, 2.07]
Gastpar 2005	100	9.8 (5.2)	94	9.3 (5.9)		12.8 %	0.50 [-1.07, 2.07]
Gastpar 2006	131	10.3 (6.4)	127	10.3 (6.4)		12.8 %	0.0 [-1.56, 1.56]
HDTSG 2002	82	12.9 (7.1)	77	10.7 (5.9)		10.2 %	2.20 [0.18, 4.22]
Schrader 2000	125	10.92 (4.27)	113	11.47 (4.04)		16.2 %	-0.55 [-1.61, 0.51]
Szegedi 2005	122	11.2 (9)	122	14.2 (8.9)		9.1 %	-3.00 [-5.25, -0.75]
van Gurp 2002	44	10.23 (8.16)	43	13.19 (7.75)		5.4 %	-2.96 [-6.30, 0.38]
Subtotal (95% CI)	**649**		**619**			**76.1 %**	**-0.38 [-1.46, 0.69]**
Heterogeneity: $Tau^2 = 1.16$; $Chi^2 = 15.40$, df = 7 (P = 0.03); $I^2 = 55\%$							
Test for overall effect: Z = 0.70 (P = 0.48)							
Total (95% CI)	**847**		**812**			**100.0 %**	**-0.34 [-1.24, 0.57]**
Heterogeneity: $Tau^2 = 1.02$; $Chi^2 = 19.21$, df = 9 (P = 0.02); $I^2 = 53\%$							
Test for overall effect: Z = 0.73 (P = 0.47)							



Review: St John's wort for major depression

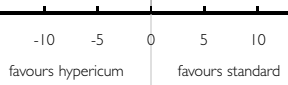
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 vs. older antidepressants						
Philipp 1999	100	9.2 (6.2)	105	10.6 (6.1)		-1.40 [-3.08, 0.28]
Vorbach 1997	98	14.4 (6.1)	88	13.4 (5.9)		1.00 [-0.73, 2.73]
Subtotal (95% CI)	**198**		**193**			**-0.21 [-2.56, 2.14]**

Heterogeneity: Tau² = 2.12; Chi² = 3.80, df = 1 (P = 0.05); I² =74%

Test for overall effect: Z = 0.17 (P = 0.86)



Review: St John's wort for major depression

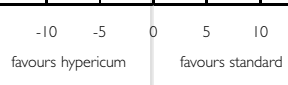
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 4 Mean HAMD (Hamilton Rating Scale for Depression) scores after 6 to 8 weeks of treatment

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Brenner 2000	13	12.7 (6.7)	15	12.5 (5.6)		0.20 [-4.41, 4.81]
Fava 2005	32	10.3 (5.9)	28	11.2 (5.8)		-0.90 [-3.87, 2.07]
Gastpar 2005	100	9.8 (5.2)	94	9.3 (5.9)		0.50 [-1.07, 2.07]
Gastpar 2006	131	10.3 (6.4)	127	10.3 (6.4)		0.0 [-1.56, 1.56]
HDTSG 2002	82	12.9 (7.1)	77	10.7 (5.9)		2.20 [0.18, 4.22]
Schrader 2000	125	10.92 (4.27)	113	11.47 (4.04)		-0.55 [-1.61, 0.51]
Szegedi 2005	122	11.2 (9)	122	14.2 (8.9)		-3.00 [-5.25, -0.75]
van Gurp 2002	44	10.23 (8.16)	43	13.19 (7.75)		-2.96 [-6.30, 0.38]
Subtotal (95% CI)	**649**		**619**			**-0.38 [-1.46, 0.69]**

Heterogeneity: Tau² = 1.16; Chi² = 15.40, df = 7 (P = 0.03); I² =55%

Test for overall effect: Z = 0.70 (P = 0.48)



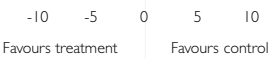
Analysis 5.5. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment

Study or subgroup	Hyericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Philipp 1999	100	-15.4 (8.1)	110	-14.2 (7.3)		10.5 %	-1.20 [-3.29, 0.89]
Subtotal (95% CI)	**100**		**110**			**10.5 %**	**-1.20 [-3.29, 0.89]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.12 (P = 0.26)							
2 vs. SSRIs							
Behnke 2002	27	-10 (5.8)	32	-12 (6.8)		5.8 %	2.00 [-1.22, 5.22]
Bjerkenstedt 2005	54	-9.9 (8.1)	54	-8.9 (8)		6.3 %	-1.00 [-4.04, 2.04]
Brenner 2000	13	-8.4 (6.5)	15	-9.1 (5.2)		3.4 %	0.70 [-3.71, 5.11]
Gastpar 2005	101	-13.7 (5.4)	97	-14 (5.5)		14.6 %	0.30 [-1.22, 1.82]
Gastpar 2006	131	-11.6 (6.3)	127	-11.4 (6.5)		14.3 %	-0.20 [-1.76, 1.36]
HDTSG 2002	113	-8.68 (7.23)	109	-10.53 (7.52)		11.4 %	1.85 [-0.09, 3.79]
Schrader 2000	125	-8.11 (5)	113	-7.25 (4.5)		17.5 %	-0.86 [-2.07, 0.35]
Szegedi 2005	122	-14.4 (8.8)	122	-11.4 (8.6)		10.0 %	-3.00 [-5.18, -0.82]
van Gurp 2002	44	-9.5 (7.1)	43	-8.2 (7.5)		6.2 %	-1.30 [-4.37, 1.77]
Subtotal (95% CI)	**730**		**712**			**89.5 %**	**-0.25 [-1.21, 0.71]**
Heterogeneity: $Tau^2 = 0.91$; $Chi^2 = 14.80$, df = 8 (P = 0.06); $I^2 = 46\%$							
Test for overall effect: Z = 0.51 (P = 0.61)							
Total (95% CI)	**830**		**822**			**100.0 %**	**-0.35 [-1.23, 0.52]**
Heterogeneity: $Tau^2 = 0.76$; $Chi^2 = 15.46$, df = 9 (P = 0.08); $I^2 = 42\%$							
Test for overall effect: Z = 0.79 (P = 0.43)							

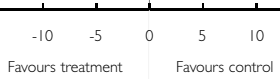


Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment

Study or subgroup	Hyericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 vs. older antidepressants						
Philipp 1999	100	-15.4 (8.1)	110	-14.2 (7.3)		-1.20 [-3.29, 0.89]
Subtotal (95% CI)	**100**		**110**			**-1.20 [-3.29, 0.89]**
Heterogeneity: not applicable						
Test for overall effect: Z = 1.12 (P = 0.26)						

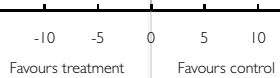


Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 5 Difference HAMD (Hamilton Rating Scale for Depression) baseline - end of treatment

Study or subgroup	Hyericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Behnke 2002	27	-10 (5.8)	32	-12 (6.8)		2.00 [-1.22, 5.22]
Bjerkenstedt 2005	54	-9.9 (8.1)	54	-8.9 (8)		-1.00 [-4.04, 2.04]
Brenner 2000	13	-8.4 (6.5)	15	-9.1 (5.2)		0.70 [-3.71, 5.11]
Gastpar 2005	101	-13.7 (5.4)	97	-14 (5.5)		0.30 [-1.22, 1.82]
Gastpar 2006	131	-11.6 (6.3)	127	-11.4 (6.5)		-0.20 [-1.76, 1.36]
HDTSG 2002	113	-8.68 (7.23)	109	-10.53 (7.52)		1.85 [-0.09, 3.79]
Schrader 2000	125	-8.11 (5)	113	-7.25 (4.5)		-0.86 [-2.07, 0.35]
Szegedi 2005	122	-14.4 (8.8)	122	-11.4 (8.6)		-3.00 [-5.18, -0.82]
van Gurp 2002	44	-9.5 (7.1)	43	-8.2 (7.5)		-1.30 [-4.37, 1.77]
Subtotal (95% CI)	**730**		**712**			**-0.25 [-1.21, 0.71]**
Heterogeneity: $Tau^2 = 0.91$; $Chi^2 = 14.80$, df = 8 (P = 0.06); $I^2 = 46\%$						
Test for overall effect: Z = 0.51 (P = 0.61)						



Analysis 5.6. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 6 MADRS after treatment.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 6 MADRS after treatment

Study or subgroup	Treatment N	Mean(SD)	Control N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Subtotal (95% CI)	**0**		**0**			**0.0 %**	**Not estimable**
Heterogeneity: not applicable							
Test for overall effect: not applicable							
2 vs. SSRIs							
Bjerkenstedt 2005	54	15.6 (10.4)	54	16.5 (9.9)		100.0 %	-0.90 [-4.73, 2.93]
Subtotal (95% CI)	**54**		**54**			**100.0 %**	**-0.90 [-4.73, 2.93]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.46 (P = 0.65)							
Total (95% CI)	**54**		**54**			**100.0 %**	**-0.90 [-4.73, 2.93]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.46 (P = 0.65)							

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 6 MADRS after treatment

Study or subgroup	Treatment N	Mean(SD)	Control N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Bjerkenstedt 2005	54	15.6 (10.4)	54	16.5 (9.9)		-0.90 [-4.73, 2.93]
Subtotal (95% CI)	**54**		**54**			**-0.90 [-4.73, 2.93]**
Heterogeneity: not applicable						
Test for overall effect: Z = 0.46 (P = 0.65)						

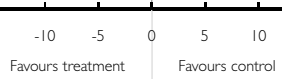
Analysis 5.7. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 7 Difference MADRS baseline - end of treatment.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 7 Difference MADRS baseline - end of treatment

Study or subgroup	Treatment N	Mean(SD)	Control N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Subtotal (95% CI)	**0**		**0**			**0.0 %**	**Not estimable**
Heterogeneity: not applicable							
Test for overall effect: not applicable							
2 vs. SSRIs							
Bjerkenstedt 2005	54	-9.9 (9.1)	54	-8.4 (8.9)		39.2 %	-1.50 [-4.89, 1.89]
Szegedi 2005	122	-16.4 (10.7)	122	-12.6 (10.6)		60.8 %	-3.80 [-6.47, -1.13]
Subtotal (95% CI)	**176**		**176**			**100.0 %**	**-2.90 [-5.10, -0.70]**
Heterogeneity: Tau² = 0.22; Chi² = 1.09, df = 1 (P = 0.30); I² =8%							
Test for overall effect: Z = 2.58 (P = 0.0099)							
Total (95% CI)	**176**		**176**			**100.0 %**	**-2.90 [-5.10, -0.70]**
Heterogeneity: Tau² = 0.22; Chi² = 1.09, df = 1 (P = 0.30); I² =8%							
Test for overall effect: Z = 2.58 (P = 0.0099)							

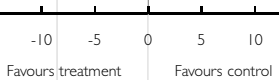


Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 7 Difference MADRS baseline - end of treatment

Study or subgroup	Treatment N	Mean(SD)	Control N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Bjerkenstedt 2005	54	-9.9 (9.1)	54	-8.4 (8.9)		-1.50 [-4.89, 1.89]
Szegedi 2005	122	-16.4 (10.7)	122	-12.6 (10.6)		-3.80 [-6.47, -1.13]
Subtotal (95% CI)	**176**		**176**			**-2.90 [-5.10, -0.70]**
Heterogeneity: Tau² = 0.22; Chi² = 1.09, df = 1 (P = 0.30); I² =8%						
Test for overall effect: Z = 2.58 (P = 0.0099)						



Analysis 5.8. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 8 Mean HAMD after treatment in studies from German-speaking countries and other studies.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 8 Mean HAMD after treatment in studies from German-speaking countries and other studies

Study or subgroup	Hypericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries							
Gastpar 2005	101	8.3 (5.5)	97	8.1 (5.6)		11.8 %	0.20 [-1.35, 1.75]
Gastpar 2006	131	10.3 (6.4)	127	10.3 (6.4)		11.7 %	0.0 [-1.56, 1.56]
Harrer 1993	44	11.2 (9.9)	42	10.5 (7.5)		4.1 %	0.70 [-3.00, 4.40]
Harrer 1999	70	7.91 (0)	79	8.11 (0)		0.0 %	Not estimable
Philipp 1999	100	9.2 (6.2)	105	10.6 (6.1)		11.0 %	-1.40 [-3.08, 0.28]
Schrader 2000	125	10.92 (4.27)	113	11.47 (4.04)		15.0 %	-0.55 [-1.61, 0.51]
Szegedi 2005	122	11.2 (9)	122	14.2 (8.9)		8.2 %	-3.00 [-5.25, -0.75]
Vorbach 1997	98	14.4 (6.1)	88	13.4 (5.9)		10.7 %	1.00 [-0.73, 2.73]
Woelk 2000	157	10.4 (0)	167	9.35 (0)		0.0 %	Not estimable
Subtotal (95% CI)	**948**		**940**			**72.3 %**	**-0.43 [-1.28, 0.41]**
Heterogeneity: $Tau^2 = 0.52$; $Chi^2 = 10.27$, df = 6 (P = 0.11); $I^2 = 42\%$							
Test for overall effect: Z = 1.00 (P = 0.32)							
2 Studies from other countries							
Behnke 2002	29	10 (0)	32	8.7 (0)		0.0 %	Not estimable
Bjerkenstedt 2005	54	15 (8.4)	54	14.9 (8.4)		5.2 %	0.10 [-3.07, 3.27]
Brenner 2000	13	12.7 (6.7)	15	12.5 (5.6)		2.8 %	0.20 [-4.41, 4.81]
Fava 2005	45	10.2 (6.6)	47	13.3 (7.3)		6.0 %	-3.10 [-5.94, -0.26]
HDTSG 2002	82	12.9 (7.1)	77	10.7 (5.9)		9.2 %	2.20 [0.18, 4.22]
van Gurp 2002	44	9.43 (8.26)	43	11.56 (8.41)		4.4 %	-2.13 [-5.63, 1.37]
Subtotal (95% CI)	**267**		**268**			**27.7 %**	**-0.44 [-2.67, 1.79]**
Heterogeneity: $Tau^2 = 3.88$; $Chi^2 = 10.54$, df = 4 (P = 0.03); $I^2 = 62\%$							
Test for overall effect: Z = 0.38 (P = 0.70)							
Total (95% CI)	**1215**		**1208**			**100.0 %**	**-0.39 [-1.23, 0.45]**
Heterogeneity: $Tau^2 = 0.94$; $Chi^2 = 21.13$, df = 11 (P = 0.03); $I^2 = 48\%$							
Test for overall effect: Z = 0.91 (P = 0.36)							

Review: St John's wort for major depression

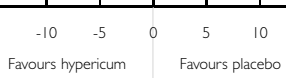
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 8 Mean HAMD after treatment in studies from German-speaking countries and other studies

Study or subgroup	Hypericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries						
Gastpar 2005	101	8.3 (5.5)	97	8.1 (5.6)		0.20 [-1.35, 1.75]
Gastpar 2006	131	10.3 (6.4)	127	10.3 (6.4)		0.0 [-1.56, 1.56]
Harrer 1993	44	11.2 (9.9)	42	10.5 (7.5)		0.70 [-3.00, 4.40]
Harrer 1999	70	7.91 (0)	79	8.11 (0)		Not estimable
Philipp 1999	100	9.2 (6.2)	105	10.6 (6.1)		-1.40 [-3.08, 0.28]
Schrader 2000	125	10.92 (4.27)	113	11.47 (4.04)		-0.55 [-1.61, 0.51]
Szegedi 2005	122	11.2 (9)	122	14.2 (8.9)		-3.00 [-5.25, -0.75]
Vorbach 1997	98	14.4 (6.1)	88	13.4 (5.9)		1.00 [-0.73, 2.73]
Woelk 2000	157	10.4 (0)	167	9.35 (0)		Not estimable
Subtotal (95% CI)	**948**		**940**			**-0.43 [-1.28, 0.41]**

Heterogeneity: $Tau^2 = 0.52$; $Chi^2 = 10.27$, df = 6 (P = 0.11); $I^2 = 42\%$

Test for overall effect: Z = 1.00 (P = 0.32)



Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 8 Mean HAMD after treatment in studies from German-speaking countries and other studies

Study or subgroup	Hypericum N	Mean(SD)	Placebo N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 Studies from other countries						
Behnke 2002	29	10 (0)	32	8.7 (0)		Not estimable
Bjerkenstedt 2005	54	15 (8.4)	54	14.9 (8.4)		0.10 [-3.07, 3.27]
Brenner 2000	13	12.7 (6.7)	15	12.5 (5.6)		0.20 [-4.41, 4.81]
Fava 2005	45	10.2 (6.6)	47	13.3 (7.3)		-3.10 [-5.94, -0.26]
HDTSG 2002	82	12.9 (7.1)	77	10.7 (5.9)		2.20 [0.18, 4.22]
van Gurp 2002	44	9.43 (8.26)	43	11.56 (8.41)		-2.13 [-5.63, 1.37]
Subtotal (95% CI)	**267**		**268**			**-0.44 [-2.67, 1.79]**

Heterogeneity: Tau² = 3.88; Chi² = 10.54, df = 4 (P = 0.03); I² =62%

Test for overall effect: Z = 0.38 (P = 0.70)

Analysis 5.9. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 9 Mean D-S (Depression Scale von Zerssen) scores after therapy.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 9 Mean D-S (Depression Scale von Zerssen) scores after therapy

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 vs. older antidepressants							
Harrer 1993	44	16.1 (15)	42	13.9 (12)		10.3 %	2.20 [-3.53, 7.93]
Vorbach 1997	98	16.5 (8.4)	88	13.6 (6.8)		70.4 %	2.90 [0.71, 5.09]
Subtotal (95% CI)	**142**		**130**			**80.6 %**	**2.81 [0.77, 4.85]**

Heterogeneity: Tau² = 0.0; Chi² = 0.05, df = 1 (P = 0.82); I² =0.0%

Test for overall effect: Z = 2.70 (P = 0.0070)

2 vs. SSRIs							
Behnke 2002	29	14.3 (10.4)	32	12.4 (8.3)		14.9 %	1.90 [-2.85, 6.65]
Brenner 2000	12	18.4 (12.1)	15	15.9 (10.5)		4.5 %	2.50 [-6.17, 11.17]
Subtotal (95% CI)	**41**		**47**			**19.4 %**	**2.04 [-2.13, 6.21]**

(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Heterogeneity: Tau² = 0.0; Chi² = 0.01, df = 1 (P = 0.91); I² =0.0%							
Test for overall effect: Z = 0.96 (P = 0.34)							
Total (95% CI)	**183**		**177**			**100.0 %**	**2.66 [0.83, 4.50]**
Heterogeneity: Tau² = 0.0; Chi² = 0.17, df = 3 (P = 0.98); I² =0.0%							
Test for overall effect: Z = 2.84 (P = 0.0045)							

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 9 Mean D-S (Depression Scale von Zerssen) scores after therapy

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
1 vs. older antidepressants						
Harrer 1993	44	16.1 (15)	42	13.9 (12)		2.20 [-3.53, 7.93]
Vorbach 1997	98	16.5 (8.4)	88	13.6 (6.8)		2.90 [0.71, 5.09]
Subtotal (95% CI)	**142**		**130**			**2.81 [0.77, 4.85]**
Heterogeneity: Tau² = 0.0; Chi² = 0.05, df = 1 (P = 0.82); I² =0.0%						
Test for overall effect: Z = 2.70 (P = 0.0070)						

Review: St John's wort for major depression

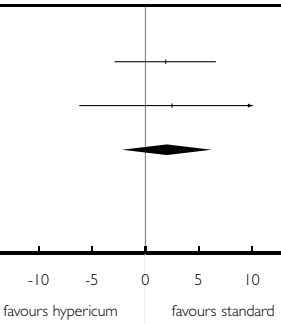
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 9 Mean D-S (Depression Scale von Zerssen) scores after therapy

Study or subgroup	hypericum N	Mean(SD)	standard N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
2 vs. SSRIs						
Behnke 2002	29	14.3 (10.4)	32	12.4 (8.3)		1.90 [-2.85, 6.65]
Brenner 2000	12	18.4 (12.1)	15	15.9 (10.5)		2.50 [-6.17, 11.17]
Subtotal (95% CI)	**41**		**47**			**2.04 [-2.13, 6.21]**

Heterogeneity: Tau² = 0.0; Chi² = 0.01, df = 1 (P = 0.91); I² =0.0%

Test for overall effect: Z = 0.96 (P = 0.34)



Analysis 5.10. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 10 Various self-rating scales.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
1 von Zerssen Depression Scale (D-S) after treatment							
Behnke 2002	29	14.3 (10.4)	32	12.4 (8.3)		5.7 %	0.20 [-0.30, 0.70]
Brenner 2000	12	18.4 (12.1)	15	15.9 (10.5)		2.9 %	0.22 [-0.55, 0.98]
Harrer 1993	44	16.1 (15)	42	13.9 (12)		7.4 %	0.16 [-0.26, 0.58]
Vorbach 1997	98	16.5 (8.4)	88	13.6 (6.8)		11.7 %	0.38 [0.09, 0.67]
Subtotal (95% CI)	**183**		**177**			**27.8 %**	**0.28 [0.07, 0.49]**
Heterogeneity: Tau² = 0.0; Chi² = 0.85, df = 3 (P = 0.84); I² =0.0%							
Test for overall effect: Z = 2.66 (P = 0.0079)							
2 Beck Depression Inventory							
van Gurp 2002	40	12 (8.2)	43	12.1 (10.1)		7.2 %	-0.01 [-0.44, 0.42]
Subtotal (95% CI)	**40**		**43**			**7.2 %**	**-0.01 [-0.44, 0.42]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.05 (P = 0.96)							
3 Beck Depression Inventory difference baseline - after treatment							
HDTSG 2002	113	-7.84 (9.67)	109	-8.75 (9.92)		12.9 %	0.09 [-0.17, 0.36]

(*Continued . . .*)

(. . . *Continued*)

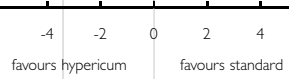
Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
Szegedi 2005	122	-10.2 (10.3)	122	-7 (9.3)		13.4 %	-0.33 [-0.58, -0.07]
Subtotal (95% CI)	**235**		**231**			**26.4 %**	**-0.12 [-0.53, 0.29]**
Heterogeneity: Tau² = 0.07; Chi² = 5.03, df = 1 (P = 0.02); I² =80%							
Test for overall effect: Z = 0.56 (P = 0.57)							
4 Zung Self Rating Depression Scale (SDS) difference baseline - after treatment							
Philipp 1999	100	-17.4 (13.6)	105	-16 (12.2)		12.5 %	-0.11 [-0.38, 0.17]
Subtotal (95% CI)	**100**		**105**			**12.5 %**	**-0.11 [-0.38, 0.17]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.77 (P = 0.44)							
5 von Zerssen Adjective Mood Scale							
Gastpar 2005	100	19.8 (13.7)	98	21.4 (16.3)		12.2 %	-0.11 [-0.38, 0.17]
Gastpar 2006	131	21.1 (11.6)	127	21.33 (11.6)		13.9 %	-0.02 [-0.26, 0.22]
Subtotal (95% CI)	**231**		**225**			**26.1 %**	**-0.06 [-0.24, 0.13]**
Heterogeneity: Tau² = 0.0; Chi² = 0.21, df = 1 (P = 0.65); I² =0.0%							
Test for overall effect: Z = 0.61 (P = 0.54)							
Total (95% CI)	**789**		**781**			**100.0 %**	**0.01 [-0.13, 0.15]**
Heterogeneity: Tau² = 0.02; Chi² = 15.91, df = 9 (P = 0.07); I² =43%							
Test for overall effect: Z = 0.17 (P = 0.86)							

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
1 von Zerssen Depression Scale (D-S) after treatment						
Behnke 2002	29	14.3 (10.4)	32	12.4 (8.3)		0.20 [-0.30, 0.70]
Brenner 2000	12	18.4 (12.1)	15	15.9 (10.5)		0.22 [-0.55, 0.98]
Harrer 1993	44	16.1 (15)	42	13.9 (12)		0.16 [-0.26, 0.58]
Vorbach 1997	98	16.5 (8.4)	88	13.6 (6.8)		0.38 [0.09, 0.67]
Subtotal (95% CI)	**183**		**177**			**0.28 [0.07, 0.49]**
Heterogeneity: Tau² = 0.0; Chi² = 0.85, df = 3 (P = 0.84); I² =0.0%						
Test for overall effect: Z = 2.66 (P = 0.0079)						



Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
2 Beck Depression Inventory						
van Gurp 2002	40	12 (8.2)	43	12.1 (10.1)		-0.01 [-0.44, 0.42]
Subtotal (95% CI)	**40**		**43**			**-0.01 [-0.44, 0.42]**

Heterogeneity: not applicable

Test for overall effect: Z = 0.05 (P = 0.96)

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
3 Beck Depression Inventory difference baseline - after treatment						
HDTSG 2002	113	-7.84 (9.67)	109	-8.75 (9.92)		0.09 [-0.17, 0.36]
Szegedi 2005	122	-10.2 (10.3)	122	-7 (9.3)		-0.33 [-0.58, -0.07]
Subtotal (95% CI)	**235**		**231**			**-0.12 [-0.53, 0.29]**

Heterogeneity: $Tau^2 = 0.07$; $Chi^2 = 5.03$, df = 1 (P = 0.02); $I^2 = 80\%$

Test for overall effect: Z = 0.56 (P = 0.57)

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
4 Zung Self Rating Depression Scale (SDS) difference baseline - after treatment						
Philipp 1999	100	-17.4 (13.6)	105	-16 (12.2)		-0.11 [-0.38, 0.17]
Subtotal (95% CI)	**100**		**105**			**-0.11 [-0.38, 0.17]**

Heterogeneity: not applicable

Test for overall effect: Z = 0.77 (P = 0.44)

Review: St John's wort for major depression

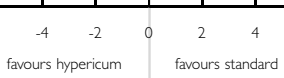
Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 10 Various self-rating scales

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
5 von Zerssen Adjective Mood Scale						
Gastpar 2005	100	19.8 (13.7)	98	21.4 (16.3)		-0.11 [-0.38, 0.17]
Gastpar 2006	131	21.1 (11.6)	127	21.33 (11.6)		-0.02 [-0.26, 0.22]
Subtotal (95% CI)	**231**		**225**			**-0.06 [-0.24, 0.13]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.21$, df = 1 (P = 0.65); $I^2 = 0.0\%$

Test for overall effect: Z = 0.61 (P = 0.54)



Analysis 5.11. Comparison 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures, Outcome 11 Various self-rating scales in studies from German-speaking countries and other countries.

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 11 Various self-rating scales in studies from German-speaking countries and other countries

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries							
Gastpar 2005	100	19.8 (13.7)	98	21.4 (16.3)		12.2 %	-0.11 [-0.38, 0.17]
Gastpar 2006	131	21.1 (11.6)	127	21.33 (11.6)		13.9 %	-0.02 [-0.26, 0.22]
Harrer 1993	44	16.1 (15)	42	13.9 (12)		7.4 %	0.16 [-0.26, 0.58]
Philipp 1999	100	-17.4 (13.6)	105	-16 (12.2)		12.5 %	-0.11 [-0.38, 0.17]
Szegedi 2005	122	-10.2 (10.3)	122	-7 (9.3)		13.4 %	-0.33 [-0.58, -0.07]
Vorbach 1997	98	16.5 (8.4)	88	13.6 (6.8)		11.7 %	0.38 [0.09, 0.67]
Subtotal (95% CI)	**595**		**582**			**71.2 %**	**-0.02 [-0.21, 0.18]**
Heterogeneity: $Tau^2 = 0.04$; $Chi^2 = 14.10$, df = 5 (P = 0.01); $I^2 = 65\%$							
Test for overall effect: Z = 0.19 (P = 0.85)							
2 Studies from other countries							
Behnke 2002	29	14.3 (10.4)	32	12.4 (8.3)		5.7 %	0.20 [-0.30, 0.70]
Brenner 2000	12	18.4 (12.1)	15	15.9 (10.5)		2.9 %	0.22 [-0.55, 0.98]
HDTSG 2002	113	-7.84 (9.67)	109	-8.75 (9.92)		12.9 %	0.09 [-0.17, 0.36]
van Gurp 2002	40	12 (8.2)	43	12.1 (10.1)		7.2 %	-0.01 [-0.44, 0.42]
Subtotal (95% CI)	**194**		**199**			**28.8 %**	**0.10 [-0.10, 0.29]**
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.50$, df = 3 (P = 0.92); $I^2 = 0.0\%$							
Test for overall effect: Z = 0.95 (P = 0.34)							
Total (95% CI)	**789**		**781**			**100.0 %**	**0.01 [-0.13, 0.15]**
Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 15.91$, df = 9 (P = 0.07); $I^2 = 43\%$							
Test for overall effect: Z = 0.17 (P = 0.86)							

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 11 Various self-rating scales in studies from German-speaking countries and other countries

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
1 Studies from German-speaking countries						
Gastpar 2005	100	19.8 (13.7)	98	21.4 (16.3)		-0.11 [-0.38, 0.17]
Gastpar 2006	131	21.1 (11.6)	127	21.33 (11.6)		-0.02 [-0.26, 0.22]
Harrer 1993	44	16.1 (15)	42	13.9 (12)		0.16 [-0.26, 0.58]
Philipp 1999	100	-17.4 (13.6)	105	-16 (12.2)		-0.11 [-0.38, 0.17]
Szegedi 2005	122	-10.2 (10.3)	122	-7 (9.3)		-0.33 [-0.58, -0.07]
Vorbach 1997	98	16.5 (8.4)	88	13.6 (6.8)		0.38 [0.09, 0.67]
Subtotal (95% CI)	**595**		**582**			**-0.02 [-0.21, 0.18]**

Heterogeneity: Tau² = 0.04; Chi² = 14.10, df = 5 (P = 0.01); I² =65%

Test for overall effect: Z = 0.19 (P = 0.85)

Review: St John's wort for major depression

Comparison: 5 Hypericum mono-preparations vs. standard antidepressants. B. Continuous measures

Outcome: 11 Various self-rating scales in studies from German-speaking countries and other countries

Study or subgroup	hypericum N	Mean(SD)	placebo N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Std. Mean Difference IV,Random,95% CI
2 Studies from other countries						
Behnke 2002	29	14.3 (10.4)	32	12.4 (8.3)		0.20 [-0.30, 0.70]
Brenner 2000	12	18.4 (12.1)	15	15.9 (10.5)		0.22 [-0.55, 0.98]
HDTSG 2002	113	-7.84 (9.67)	109	-8.75 (9.92)		0.09 [-0.17, 0.36]
van Gurp 2002	40	12 (8.2)	43	12.1 (10.1)		-0.01 [-0.44, 0.42]
Subtotal (95% CI)	**194**		**199**			**0.10 [-0.10, 0.29]**

Heterogeneity: Tau² = 0.0; Chi² = 0.50, df = 3 (P = 0.92); I² =0.0%

Test for overall effect: Z = 0.95 (P = 0.34)

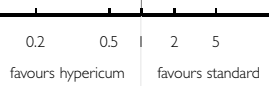
Analysis 6.1. Comparison 6 Safety - Hypericum mono-preparations vs. standard antidepressants, Outcome 1 Number of patients discontinuing treatment/dropping out due to adverse/side effects.

Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 1 Number of patients discontinuing treatment/dropping out due to adverse/side effects

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
1 vs. older antidepressants					
Harrer 1993	0/51	2/51		1.4 %	0.19 [0.01, 4.11]
Philipp 1999	0/106	1/110		1.3 %	0.34 [0.01, 8.51]
Vorbach 1997	1/107	8/102		3.1 %	0.11 [0.01, 0.90]
Wheatley 1997	7/87	13/78		14.2 %	0.44 [0.16, 1.16]
Woelk 2000	4/157	26/167		11.7 %	0.14 [0.05, 0.42]
Subtotal (95% CI)	**508**	**508**		**31.7 %**	**0.24 [0.13, 0.46]**
Total events: 12 (hypericum), 50 (standard)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 3.03, df = 4 (P = 0.55); I^2 =0.0%					
Test for overall effect: Z = 4.27 (P = 0.000020)					
2 vs. SSRIs					
Behnke 2002	2/35	2/35		3.3 %	1.00 [0.13, 7.53]
Bjerkenstedt 2005	4/57	4/56		6.5 %	0.98 [0.23, 4.13]
Brenner 2000	2/15	2/15		3.1 %	1.00 [0.12, 8.21]
Fava 2005	0/45	2/47		1.4 %	0.20 [0.01, 4.28]
Gastpar 2005	5/123	10/118		11.1 %	0.46 [0.15, 1.38]
Gastpar 2006	4/131	11/127		9.9 %	0.33 [0.10, 1.07]
Harrer 1999	6/77	8/84		11.0 %	0.80 [0.27, 2.43]
HDTSG 2002	2/113	5/111		4.9 %	0.38 [0.07, 2.01]
Schrader 2000	0/125	1/114		1.3 %	0.30 [0.01, 7.47]
Szegedi 2005	4/125	8/126		9.0 %	0.49 [0.14, 1.66]
van Gurp 2002	3/45	7/45		6.7 %	0.39 [0.09, 1.61]
Subtotal (95% CI)	**891**	**878**		**68.3 %**	**0.53 [0.34, 0.83]**
Total events: 32 (hypericum), 60 (standard)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 3.52, df = 10 (P = 0.97); I^2 =0.0%					
Test for overall effect: Z = 2.80 (P = 0.0050)					
Total (95% CI)	**1399**	**1386**		**100.0 %**	**0.41 [0.29, 0.60]**
Total events: 44 (hypericum), 110 (standard)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 10.42, df = 15 (P = 0.79); I^2 =0.0%					
Test for overall effect: Z = 4.72 (P < 0.00001)					



Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 1 Number of patients discontinuing treatment/dropping out due to adverse/side effects

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
1 vs. older antidepressants				
Harrer 1993	0/51	2/51		0.19 [0.01, 4.11]
Philipp 1999	0/106	1/110		0.34 [0.01, 8.51]
Vorbach 1997	1/107	8/102		0.11 [0.01, 0.90]
Wheatley 1997	7/87	13/78		0.44 [0.16, 1.16]
Woelk 2000	4/157	26/167		0.14 [0.05, 0.42]
Subtotal (95% CI)	**508**	**508**		**0.24 [0.13, 0.46]**
Total events: 12 (hypericum), 50 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.03$, df = 4 (P = 0.55); $I^2 = 0.0\%$				
Test for overall effect: Z = 4.27 (P = 0.000020)				

0.2 0.5 1 2 5

favours hypericum favours standard

Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 1 Number of patients discontinuing treatment/dropping out due to adverse/side effects

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
2 vs. SSRIs				
Behnke 2002	2/35	2/35		1.00 [0.13, 7.53]
Bjerkenstedt 2005	4/57	4/56		0.98 [0.23, 4.13]
Brenner 2000	2/15	2/15		1.00 [0.12, 8.21]
Fava 2005	0/45	2/47		0.20 [0.01, 4.28]
Gastpar 2005	5/123	10/118		0.46 [0.15, 1.38]
Gastpar 2006	4/131	11/127		0.33 [0.10, 1.07]
Harrer 1999	6/77	8/84		0.80 [0.27, 2.43]
HDTSG 2002	2/113	5/111		0.38 [0.07, 2.01]
Schrader 2000	0/125	1/114		0.30 [0.01, 7.47]

0.2 0.5 1 2 5

favours hypericum favours standard

(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
Szegedi 2005	4/125	8/126		0.49 [0.14, 1.66]
van Gurp 2002	3/45	7/45		0.39 [0.09, 1.61]
Subtotal (95% CI)	**891**	**878**		**0.53 [0.34, 0.83]**
Total events: 32 (hypericum), 60 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.52$, df = 10 (P = 0.97); $I^2 = 0.0\%$				
Test for overall effect: Z = 2.80 (P = 0.0050)				

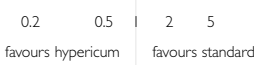
Analysis 6.2. Comparison 6 Safety - Hypericum mono-preparations vs. standard antidepressants, Outcome 2 Number of patients dropping out.

Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 2 Number of patients dropping out

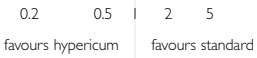
Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
1 vs. older antidperessants					
Harrer 1993	7/51	9/51		4.0 %	0.74 [0.25, 2.17]
Philipp 1999	13/106	11/110		6.3 %	1.26 [0.54, 2.95]
Vorbach 1997	9/107	14/102		5.8 %	0.58 [0.24, 1.40]
Wheatley 1997	21/87	24/78		9.7 %	0.72 [0.36, 1.42]
Woelk 2000	15/157	32/167		10.6 %	0.45 [0.23, 0.86]
Subtotal (95% CI)	**508**	**508**		**36.4 %**	**0.67 [0.47, 0.95]**
Total events: 65 (hypericum), 90 (standard)					
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.76$, df = 4 (P = 0.44); $I^2 = 0.0\%$					
Test for overall effect: Z = 2.24 (P = 0.025)					
2 vs. SSRIs					
Behnke 2002	6/35	3/35		2.1 %	2.21 [0.51, 9.64]
Bjerkenstedt 2005	9/57	8/56		4.3 %	1.13 [0.40, 3.16]
Brenner 2000	7/15	3/15		1.7 %	3.50 [0.69, 17.71]
Fava 2005	18/45	23/47		6.7 %	0.70 [0.30, 1.59]
Gastpar 2005	17/123	19/118		9.1 %	0.84 [0.41, 1.70]



(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Gastpar 2006	6/131	6/127		3.4 %	0.97 [0.30, 3.08]
Harrer 1999	8/77	16/84		5.5 %	0.49 [0.20, 1.23]
HDTSG 2002	31/113	32/111		13.5 %	0.93 [0.52, 1.67]
Schrader 2000	1/125	1/114		0.6 %	0.91 [0.06, 14.74]
Szegedi 2005	17/125	29/126		10.5 %	0.53 [0.27, 1.02]
van Gurp 2002	16/45	17/45		6.2 %	0.91 [0.39, 2.14]
Subtotal (95% CI)	**891**	**878**		**63.6 %**	**0.83 [0.63, 1.08]**
Total events: 136 (hypericum), 157 (standard)					
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 8.59$, df = 10 (P = 0.57); $I^2 = 0.0\%$					
Test for overall effect: Z = 1.39 (P = 0.17)					
Total (95% CI)	**1399**	**1386**		**100.0 %**	**0.77 [0.62, 0.95]**
Total events: 201 (hypericum), 247 (standard)					
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 13.25$, df = 15 (P = 0.58); $I^2 = 0.0\%$					
Test for overall effect: Z = 2.46 (P = 0.014)					



Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 2 Number of patients dropping out

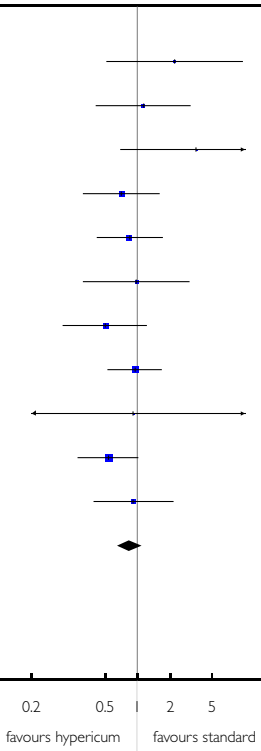
Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
1 vs. older antidperessants				
Harrer 1993	7/51	9/51		0.74 [0.25, 2.17]
Philipp 1999	13/106	11/110		1.26 [0.54, 2.95]
Vorbach 1997	9/107	14/102		0.58 [0.24, 1.40]
Wheatley 1997	21/87	24/78		0.72 [0.36, 1.42]
Woelk 2000	15/157	32/167		0.45 [0.23, 0.86]
Subtotal (95% CI)	**508**	**508**		**0.67 [0.47, 0.95]**
Total events: 65 (hypericum), 90 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.76$, df = 4 (P = 0.44); $I^2 = 0.0\%$				
Test for overall effect: Z = 2.24 (P = 0.025)				

Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 2 Number of patients dropping out

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
2 vs. SSRIs				
Behnke 2002	6/35	3/35		2.21 [0.51, 9.64]
Bjerkenstedt 2005	9/57	8/56		1.13 [0.40, 3.16]
Brenner 2000	7/15	3/15		3.50 [0.69, 17.71]
Fava 2005	18/45	23/47		0.70 [0.30, 1.59]
Gastpar 2005	17/123	19/118		0.84 [0.41, 1.70]
Gastpar 2006	6/131	6/127		0.97 [0.30, 3.08]
Harrer 1999	8/77	16/84		0.49 [0.20, 1.23]
HDTSG 2002	31/113	32/111		0.93 [0.52, 1.67]
Schrader 2000	1/125	1/114		0.91 [0.06, 14.74]
Szegedi 2005	17/125	29/126		0.53 [0.27, 1.02]
van Gurp 2002	16/45	17/45		0.91 [0.39, 2.14]
Subtotal (95% CI)	**891**	**878**		**0.83 [0.63, 1.08]**
Total events: 136 (hypericum), 157 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 8.59$, df = 10 (P = 0.57); $I^2 = 0.0\%$				
Test for overall effect: Z = 1.39 (P = 0.17)				



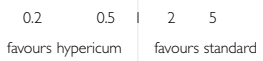
Analysis 6.3. Comparison 6 Safety - Hypericum mono-preparations vs. standard antidepressants, Outcome 3 Number of patients reporting adverse effects.

Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 3 Number of patients reporting adverse effects

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
1 vs. older antidepressants					
Harrer 1993	13/51	18/51		5.7 %	0.63 [0.27, 1.47]
Philipp 1999	23/106	51/110		8.0 %	0.32 [0.18, 0.58]
Vorbach 1997	25/107	42/102		7.9 %	0.44 [0.24, 0.79]
Wheatley 1997	32/87	50/78		7.6 %	0.33 [0.17, 0.62]
Woelk 2000	62/157	105/167		9.5 %	0.39 [0.25, 0.60]
Subtotal (95% CI)	**508**	**508**		**38.7 %**	**0.39 [0.30, 0.50]**
Total events: 155 (hypericum), 266 (standard)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 2.05, df = 4 (P = 0.73); I^2 =0.0%					
Test for overall effect: Z = 7.07 (P < 0.00001)					
2 vs. SSRIs					
Behnke 2002	22/35	20/35		5.0 %	1.27 [0.49, 3.31]
Bjerkenstedt 2005	20/57	31/56		6.5 %	0.44 [0.20, 0.93]
Gastpar 2005	74/123	60/118		8.8 %	1.46 [0.88, 2.43]
Gastpar 2006	39/131	53/127		8.8 %	0.59 [0.35, 0.99]
Harrer 1999	12/77	17/84		6.0 %	0.73 [0.32, 1.64]
HDTSG 2002	100/113	103/111		5.2 %	0.60 [0.24, 1.50]
Schrader 2000	18/125	28/114		7.4 %	0.52 [0.27, 1.00]
Szegedi 2005	69/125	96/126		8.5 %	0.39 [0.22, 0.66]
van Gurp 2002	34/45	32/45		5.1 %	1.26 [0.49, 3.20]
Subtotal (95% CI)	**831**	**816**		**61.3 %**	**0.70 [0.49, 1.00]**
Total events: 388 (hypericum), 440 (standard)					
Heterogeneity: Tau^2 = 0.16; Chi^2 = 18.48, df = 8 (P = 0.02); I^2 =57%					
Test for overall effect: Z = 1.98 (P = 0.048)					
Total (95% CI)	**1339**	**1324**		**100.0 %**	**0.56 [0.43, 0.74]**
Total events: 543 (hypericum), 706 (standard)					
Heterogeneity: Tau^2 = 0.15; Chi^2 = 31.44, df = 13 (P = 0.003); I^2 =59%					
Test for overall effect: Z = 4.14 (P = 0.000035)					



Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 3 Number of patients reporting adverse effects

Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
1 vs. older antidepressants				
Harrer 1993	13/51	18/51		0.63 [0.27, 1.47]
Philipp 1999	23/106	51/110		0.32 [0.18, 0.58]
Vorbach 1997	25/107	42/102		0.44 [0.24, 0.79]
Wheatley 1997	32/87	50/78		0.33 [0.17, 0.62]
Woelk 2000	62/157	105/167		0.39 [0.25, 0.60]
Subtotal (95% CI)	**508**	**508**		**0.39 [0.30, 0.50]**
Total events: 155 (hypericum), 266 (standard)				
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 2.05$, df = 4 (P = 0.73); $I^2 = 0.0\%$				
Test for overall effect: Z = 7.07 (P < 0.00001)				

0.2 0.5 1 2 5

favours hypericum favours standard

Review: St John's wort for major depression

Comparison: 6 Safety - Hypericum mono-preparations vs. standard antidepressants

Outcome: 3 Number of patients reporting adverse effects

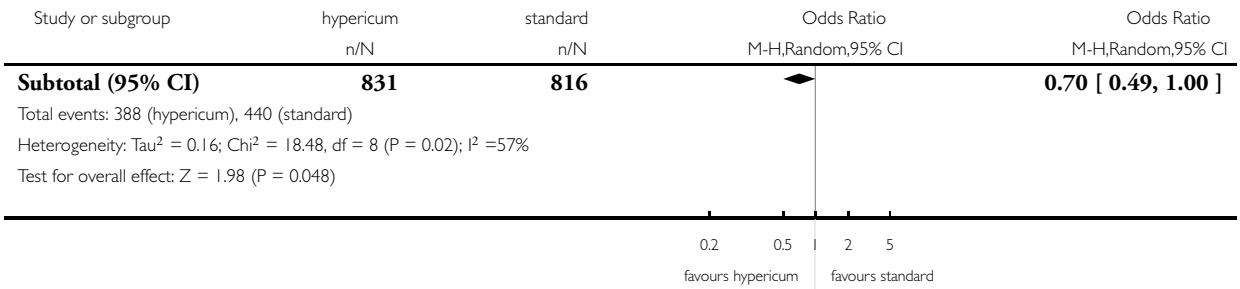
Study or subgroup	hypericum n/N	standard n/N	Odds Ratio M-H,Random,95% CI	Odds Ratio M-H,Random,95% CI
2 vs. SSRIs				
Behnke 2002	22/35	20/35		1.27 [0.49, 3.31]
Bjerkenstedt 2005	20/57	31/56		0.44 [0.20, 0.93]
Gastpar 2005	74/123	60/118		1.46 [0.88, 2.43]
Gastpar 2006	39/131	53/127		0.59 [0.35, 0.99]
Harrer 1999	12/77	17/84		0.73 [0.32, 1.64]
HDTSG 2002	100/113	103/111		0.60 [0.24, 1.50]
Schrader 2000	18/125	28/114		0.52 [0.27, 1.00]
Szegedi 2005	69/125	96/126		0.39 [0.22, 0.66]
van Gurp 2002	34/45	32/45		1.26 [0.49, 3.20]

0.2 0.5 1 2 5

favours hypericum favours standard

(*Continued . . .*)

(. . . *Continued*)



APPENDICES

Appendix 1. Summary of meta-regression analyses

Table 1: Univariable meta-regression analysis of response rate ratio (relative risk of response) in comparison to placebo

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained (R^2)	Interpretation
Country	0.53	0.20 to 0.86	.002	0.29	Studies form German-speaking countries result statistically significant in 0.53 higher effect size estimates on average than studies form non-German-speaking countries.
Precision	-0.11	-0.21 to -0.01	.032	0.16	Studies with one unit increase in precision (1/SE) result statistically significant in 0.11 lower effect size estimates on average.
Baseline HAMD	-0.12	-0.23 to 0	.048	0.16	Studies with one point increase in mean baseline HAMD severity result statistically significant in 0.12 lower effect size

(*Continued*)

estimates on average.

Table 2: Multiple meta-regression analysis of response rate ratio (relative risk of response) in comparison to placebo

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained ($Beta^2$)	Interpretation
Country	0.32	0.02 to 0.61	.035	0.13	Studies form German-speaking countries result statistically significant in 0.32 higher effect size estimates on average than studies form non-German-speaking countries.
Precision	-0.09	-0.18 to -0.02	.017	0.17	Studies with one unit increase in precision (1/SE) result statistically significant in 0.09 lower effect size estimates on average.
Baseline HAMD	-0.07	-0.16 to 0.03	.161	0.06	No statistically significant association was found.

Table 3: Univariable meta-regression analysis of response rate ratio (relative risk of response) in comparison to standard antidepressants

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained (R^2)	Interpretation
Country	0.18	0.01 to 0.36	.037	0.23	Studies form German-speaking countries result statistically significant in 0.18 higher effect size estimates on average than studies form non-German-speaking countries.
Precision	0.03	0 to 0.06	.065	0.18	No statistically significant association was found.
Baseline HAMD	0.01	-0.02 to 0.04	.546	0.02	No statistically significant association was found.

Table 4: Multiple meta-regression analysis of response rate ratio (relative risk of response) in comparison to standard antidepressants

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained ($Beta^2$)	Interpretation
Country	0.14	-0.10 to 0.38	.244	0.17	No statistically significant association was found.
Precision	0.01	-0.04 to 0.05	.832	0.01	No statistically significant association was found.
Baseline HAMD	0.01	-0.02 to 0.04	.704	0.01	No statistically significant association was found.

Table 5: Univariable meta-regression analysis of mean difference in post-treatment HAMD scores in placebo comparisons

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained (R^2)	Interpretation
Country	3.80	1.95 to 5.66	<.001	0.54	Studies form German-speaking countries result statistically significant in 3.80 HAMD points higher effect size estimates on average than studies form non-German-speaking countries.
Precision	3.02	-2.10 to 8.15	.247	0.09	No statistically significant association was found.

(*Continued*)

Baseline HAMD	-0.57	-1.27 to 0.14	.117	0.15	No statistically

Table 6: Multiple meta-regression analysis of mean difference in post-treatment HAMD scores in placebo comparisons

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained ($Beta^2$)	Interpretation
Country	3.52	1.70 to 5.33	<.001	0.46	Studies form German-speaking countries result statistically significant in 3.52 HAMD points higher effect size estimates on average than studies form non-German-speaking countries.
Precision	2.43	-1.14 to 5.99	.183	0.05	No statistically significant association was found.
Baseline HAMD	-0.24	-0.76 to 0.29	.378	0.03	No statistically significant association was found.

Table 7: Univariable meta-regression analysis of mean difference in post-treatment HAMD scores in standard antidepressant comparisons

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained (R^2)	Interpretation
Country	0.02	-1.64 to 1.68	.982	0	No statistically significant association was found.
Precision	0.04	-1.71 to 1.79	.966	0	No statistically significant association was found.
Baseline HAMD	-0.24	-0.48 to 0	.045	0.26	Studies with one point increase in mean baseline HAMD severity result statistically significant in 0.24 lower effect size estimates on average.

Table 8: Multiple meta-regression analysis of mean difference in post-treatment HAMD scores in standard antidepressant comparisons

Explanatory variable	Coefficient (B)	95% confidence interval	Statistical significance (p)	Proportion of heterogeneity explained ($Beta^2$)	Interpretation
Country	1.34	-0.76 to 3.45	.211	0.22	No statistically significant association was found.
Precision	-2.11	-4.62 to 0.40	.099	0.60	No statistically significant association was found.
Baseline HAMD	-0.45	-0.79 to 0.11	.010	0.88	Studies with one point increase in mean baseline HAMD severity result statistically significant in 0.45 lower effect size estimates on average.

WHAT'S NEW

Last assessed as up-to-date: 7 August 2008

Date	Event	Description
15 July 2008	New citation required and conclusions have changed	1) The title of the review has been changed from 'St. John's wort for depression' to 'St. John's wort for major depression'. This reflects that inclusion has been now limited to trials in patients suffering from major depression only. 2) The modification of selection criteria resulted in the exclusion of 16 of 37 studies included in the previous version. Eight new trials have been included. The review now covers 29 trials with 18 comparisons of a hypericum extract with a placebo and 17 with a standard antidepressant (six three-armed trials). 3) Compared to the previous version our conclusions are now slightly more favourable, as modest effects over placebo have also been shown in several large trials, and as side effects appear to be less frequent compared to both older antidepressants and selective serotonin reuptake inhibitors. 4) The list of authors has been amended.
23 May 2008	Amended	Converted to new review format.

HISTORY

Protocol first published: Issue 3, 1997

Review first published: Issue 4, 1998

Date	Event	Description
12 December 2007	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

Study concept and design: Linde, Berner, Kriston

Acquisition of data: Linde, Berner, Kriston

Analysis and interpretation of data: Linde, Berner, Kriston

Drafting and revision of the manuscript: Linde, Berner, Kriston

Statistical expertise: Kriston

Study coordination: Linde

DECLARATIONS OF INTEREST

Michael Berner has received a grant for research on hypericum in the past and has received fees for speaking on conferences from a manufacturer (Schwabe). Klaus Linde once received reimbursement for travel expenses for speaking on a meeting organised by a manufacturer (Schwabe). Levente Kriston once received reimbursement for travel expenses (Schwabe).

SOURCES OF SUPPORT

Internal sources

- Centre for Complementary Medicine Research, Department of Internal Medicine II, Technische Universität München, Germany.
- Department of Psychiatry and Psychotherapy, University of Freiburg, Germany.

External sources

- No sources of support supplied

INDEX TERMS

Medical Subject Headings (MeSH)

Depressive Disorder [*drug therapy]; *Hypericum; *Phytotherapy; Randomized Controlled Trials as Topic

MeSH check words

Adult; Humans