

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

A Szegedi, R Kohnen, A Dienel, M Kieser

Abstract

Objective To investigate the efficacy of hypericum extract (St John's wort) compared with paroxetine in patients with moderate to severe major depression.

Design Randomised double blind, double dummy, reference controlled, multicentre non-inferiority trial.

Setting 21 psychiatric primary care practices in Germany.

Participants 251 adult outpatients with acute major depression with total score ≥ 22 on the 17 item Hamilton depression scale.

Interventions 900 mg/day hypericum extract WS 5570 three times a day or 20 mg paroxetine once a day for six weeks. In initial non-responders doses were increased to 1800 mg/day hypericum or 40 mg/day paroxetine after two weeks.

Main outcome measures Change in score on Hamilton depression scale from baseline to day 42 (primary outcome). Secondary measures were change in scores on Montgomery-Åsberg depression rating scale, clinical global impressions, and Beck depression inventory.

Results The Hamilton depression total score decreased by mean 14.4 (SD 8.8) points, corresponding to 56.6% (SD 34.3%) of the baseline value, in the hypericum group and by 11.4 (SD 8.6) points (44.8% (SD 33.5%) of baseline value) in the paroxetine group (intention to treat analysis; similar results were observed in the per protocol analysis). The intention to treat analysis (lower one sided 97.5% confidence limit 1.5 points for the difference hypericum minus paroxetine) and the per protocol analysis (lower confidence limit 0.7 points) showed non-inferiority of hypericum and statistical superiority over paroxetine. The lower limits in both cases exceeded the pre-specified non-inferiority margin of -2.5 points and the superiority margin of 0. The incidence of adverse events was 0.035 and 0.060 events per day of exposure for hypericum and paroxetine, respectively.

Conclusions In the treatment of moderate to severe major depression, hypericum extract WS 5570 is at least as effective as paroxetine and is better tolerated.

Introduction

Extract of *Hypericum perforatum* (St John's wort) is more effective than placebo in the treatment of mild to moderate major depression¹ and as effective as several tricyclic antidepressants²⁻⁵ or fluoxetine.⁶ In patients with more severe depression, however, the antidepressant efficacy of hypericum extract is disputed. In a comparison of 1800 mg/day hypericum extract (LI 160) and 150 mg/day imipramine the effect of both drugs was comparable during six weeks of acute treatment.⁷ That study, however, was not sufficiently powered to demonstrate non-inferiority of the herbal extract.

We compared the efficacy and safety of hypericum extract WS 5570 with paroxetine in patients with moderate to severe depression over a six week acute phase after which responders undergo four months of prophylactic continuation treatment (to prevent relapse or recurrence, or both).

Methods

Protocol, design, and objectives

This double blind, double dummy, randomised trial examined the efficacy of hypericum extract compared with paroxetine in the acute treatment of moderate to severe major depression. After a screening examination participants underwent a single blind placebo run-in phase of three to seven days, after which we randomised those still meeting the selection criteria to six weeks of double blind treatment with hypericum extract or paroxetine.

Participants

We recruited male and female outpatients in 21 psychiatric primary care centres in Germany. All participants were 18-70 years old and had single or recurrent moderate or severe episodes of unipolar major depression without psychotic features persisting for two weeks to a year.

Charité-
Universitätsmedizin
Berlin, Campus
Benjamin Franklin,
Department of
Psychiatry and
Psychotherapy,
Eschenallee 3,
14050 Berlin,
Germany
A Szegedi
managing senior
physician

Institute for Medical
Research
Management and
Biometrics GmbH,
Scheurlstraße 21,
90478 Nürnberg,
Germany

R Kohnen
head of scientific
affairs

Dr Willmar
Schwabe
Pharmaceuticals,
PO Box 410925,
76209 Karlsruhe,
Germany

A Dienel
head of clinical trials
department

M Kieser
head of biometry
department

Correspondence to:
M Kieser
meinhard.kieser@
schwabe.de

BMJ 2005;330:503-6



This is the abridged version of an article that was posted on
bmj.com on 11 February 2005: <http://bmj.com/cgi/doi/10.1136/bmj.38356.655266.82>

We excluded anyone with a decrease in total depression score of $\geq 25\%$ during the run-in, or with a diagnosis of schizophrenia, acute anxiety disorder, adjustment disorder, depressive disorder of any type not stated above, bipolar disorder, organic mental disorder, acute post-traumatic stress disorder, or substance abuse disorder. We also excluded patients with increased risk of suicide, who had previously attempted suicide, or who had not responded to more than one adequate treatment in the present episode. Participants were not allowed to take other psychotropic medication and psychotherapy during the study.

Interventions and blinding

We used hypericum extract WS 5570 (Dr Willmar Schwabe Pharmaceuticals, Karlsruhe, Germany). The coated tablets contained 300 mg or 600 mg of the extract. Paroxetine was supplied in tablets of 20 mg packed in capsules containing one or two tablets. For each drug an identically matched placebo was available.

During the six weeks of randomised treatment patients allocated to hypericum always took three doses of 300 mg/day hypericum or one dose of 20 mg/day paroxetine. For patients whose total depression score had not decreased by at least 20% after two weeks we increased treatment to three doses of 600 mg/day hypericum or one dose of 40 mg/day paroxetine.

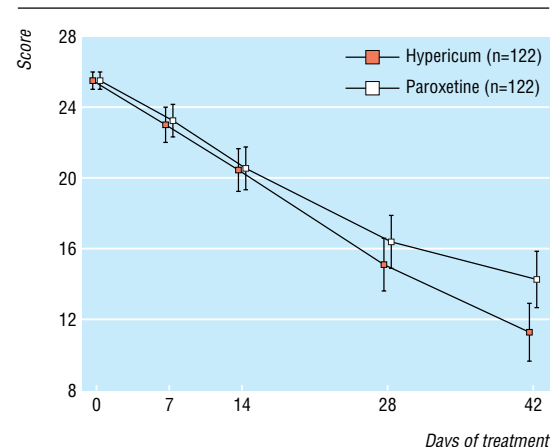
Outcomes

We assessed efficacy and safety at screening, baseline, and at the end of the first, second, fourth, and sixth weeks. The primary outcome measure was the absolute decrease of the Hamilton total depression score between baseline and week six. Secondary outcome measures included the Montgomery-Åsberg depression rating scale, the clinical global impressions, and the Beck depression inventory.

Statistical methods

We considered that hypericum would not be relevantly inferior to paroxetine if the true decrease in total depression score (primary outcome measure) for hypericum was not more than 2.5 points⁸ smaller than for paroxetine ($\delta = -2.5$).

We reserved the option of testing for superiority after establishing non-inferiority of hypericum. The primary analysis was based on the intention to treat



Total Hamilton depression scores over time (intention to treat analysis, means and 95% confidence intervals)

analysis. We also performed a per protocol analysis to demonstrate robustness of the trial result to the choice of the analysis set.⁹ For the Hamilton total score, we defined response as a decrease in total score of $\geq 50\%$ from baseline and remission as a score ≤ 10 points at week six.

Results

Participants

Between May 2000 and July 2003, we randomised 251 patients (125 to hypericum and 126 to paroxetine). Baseline demographic and clinical measures were comparable in both groups (table 1). Mean age and average duration of the current episode, however, were higher in the hypericum group. In each group more than half of the patients had a total depression score ≥ 25 and were thus severely depressed.¹⁰

Investigational treatment

After two weeks of randomised treatment, 69/122 patients in the hypericum group (57%) and 58/122 in the paroxetine group (48%) were switched to the higher doses. Between baseline and day 42 Hamilton total scores decreased by an average of 14.4 (SD 8.8) points (corresponding to 57% (SD 34%) of the baseline value) for hypericum and by 11.4 (SD 8.6) points (45% (SD 34%)) for paroxetine (lower one sided repeated 97.5% confidence limit for the difference hypericum-paroxetine was 1.5 points) (figure).

At the end of the acute treatment phase 86/122 patients (71%) in the hypericum group and 73/122 (60%) in the paroxetine group responded to treatment ($P=0.08$), and 61/122 (50%) and 43/122 patients (35%) showed remission ($P=0.02$). For all standardised psychiatric scales we found differences between treatment groups in favour of hypericum (table 2).

Safety and tolerability

During the acute treatment phase 69/125 patients randomised to hypericum (55%) reported 172 adverse events and 96/126 treated with paroxetine (76%) reported 269 adverse events. The incidences were 0.035 adverse events per day of exposure (0.029 at 900 mg/day and 0.039 at 1800 mg/day) for hypericum and 0.060 (0.062 at 20 mg/day and 0.059 at 40 mg/day) for paroxetine. Based on the rate ratio, the incidence of

Table 1 Demographic and clinical characteristics at baseline (intention to treat analysis; figures are means (SD); medians unless stated otherwise)

	Hypericum (n=122)	Paroxetine (n=122)
No (%) of women	85 (70)	83 (68)
Age (years)	49.0 (11.0); 51.5	45.5 (11.5); 48.0
No (%) with recurrent depression	50 (41)	49 (40)
Duration of current episode (days)	160 (109); 148	127 (81); 106
HAMD total score*	25.5 (2.7); 25.0	25.5 (2.9); 25.0
No (%) with HAMD total score ≥ 25	69 (57)	67 (55)
MADRS total score†	29.9 (5.0); 29.0	29.4 (4.9); 29.0
Beck depression inventory‡	26.3 (8.5); 26.0	25.6 (8.0); 24.5
No (%) markedly or severely ill§	87 (71)	84 (69)

HAMD=Hamilton depression scale; MADRS=Montgomery-Åsberg depression rating scale.

*Theoretical range 0–52.

†Theoretical range 0–60.

‡Theoretical range 0–63; 119 in hypericum group, 120 in paroxetine group.

§According to clinical global impressions score.

Table 2 Secondary measures (intention to treat analysis; figures are numbers (percentages) unless stated otherwise)

	Hypericum (n=122)	Paroxetine (n=122)	Difference (hypericum minus paroxetine) (95% CI), P value
Change from baseline to day 42:			
MADRS (mean (SD); median)	16.4 (10.7); 17.0	12.6 (10.6); 14.0	3.8 (1.1 to 6.5), 0.01*
BDI (mean (SD); median)†	10.2 (10.3); 9.0	7.0 (9.3); 5.5	3.2 (0.7 to 5.7), 0.01*
Scores by day 42:			
Clinical global impressions:			
Item 1 improved by ≥2 categories	71 (58)	52 (43)	16 (3 to 28), 0.02‡
Item 2 much or very much improved	83 (68)	70 (57)	11 (-1 to 23), 0.09‡
Item 3 marked therapeutic effect	49 (40)	36 (30)	11 (-1 to 23), 0.08‡
Global efficacy self rating very good or good	65 (53)	55 (45)	8 (-4 to 21), 0.20‡

MADRS=Montgomery-Åsberg depression rating scale; BDI=Beck depression inventory.

*t test for difference (calculated for pooled data from both study stages).

†119 in hypericum group, 120 in paroxetine group.

‡χ² test for difference (calculated for pooled data from both study stages).

adverse events in the paroxetine group was 1.72 (95% confidence interval 1.42 to 2.10) of the rate observed for hypericum. The highest incidence was found for gastrointestinal disorders (59 events in 42 patients in the hypericum group and 106 events in 67 patients in the paroxetine group), followed by nervous system disorders (35 events in 29 patients and 61 events in 43 patients, respectively). Table 3 shows adverse events that occurred in at least 10 patients in one group. Two serious adverse events occurred in the hypericum group (psychic decompensation attributable to social problems; hypertensive crisis); both were thought to be unrelated to hypericum—that is, a cause other than the administration of hypericum was evident.

Discussion

Strengths and weaknesses

Non-inferiority trials of hypericum extract against synthetic antidepressants have been criticised for using doses mostly in the lower therapeutic range of the active comparators.¹¹ Our trial included a mandatory dose increase in patients with insufficient response after two weeks of treatment. For paroxetine, 40 mg/day correspond to the established use of the drug in clinical trials and daily practice.¹² The trial's assay sensitivity is supported by the observed treatment effect for paroxetine which was in line with previously published data from trials against placebo and synthetic antidepressants.¹³ Another indicator of a pharmacological effect is that in both study groups a (single blind) dose increase in initial non-responders was followed by a substantial decrease in depression score that was comparable with the effect observed in those patients who were adequately treated with the initial (lower) dose. A placebo control could not be

Table 3 Adverse events that occurred in at least 10 patients in one group (safety analysis set; figures are numbers (percentages) of patients)

	Hypericum (n=125)	Paroxetine (n=126)
Upper abdominal pain	12 (9.6)	9 (7.1)
Diarrhoea	12 (9.6)	23 (18.3)
Dry mouth	16 (12.8)	35 (27.8)
Nausea	9 (7.2)	21 (16.7)
Fatigue	14 (11.2)	16 (12.7)
Dizziness	9 (7.2)	24 (19.1)
Headache	13 (10.4)	14 (11.1)
Sleep disorder	5 (4.0)	10 (7.9)
Increased sweating	9 (7.2)	13 (10.3)

used in this group of predominantly severely depressed patients for ethical reasons, particularly as comedication with benzodiazepines was not permitted. For the same reason we had to refrain from including patients at high risk of suicide. As we did not actually withdraw any patient because of increased risk of suicide, however, this restriction does not adversely affect the external validity of our data.

Implications for clinicians

Our results support the use of hypericum extract WS 5570 as an alternative to standard antidepressants in moderate to severe depression, especially as it is well tolerated.⁷ As in any effective antidepressant, potential interactions with other drugs deserve clinical attention.⁷

The convincing results for hypericum extract WS 5570 observed in this trial deserve independent confirmation by other research. We are assessing efficacy in long term treatment, for which the drug can be an interesting option because of its favourable ratio between efficacy and tolerability, in the ongoing continuation phase.

We thank the investigators and patients of the study, St Klement for project management, T Konstantinowicz for the data analysis, T Utz for project assistance, and A Völp for his assistance in the preparation of the manuscript.

Contributors: See bmj.com

Funding: Dr Willmar Schwabe Pharmaceuticals, manufacturer of WS 5570.

Competing interests: AS has received consultancy fees from Dr Willmar Schwabe Pharmaceuticals. RK is head of a contract research organisation (IMEREM), which is engaged in several

What is already known on this topic

Hypericum extract is effective in the acute treatment of patients with mild to moderate depression

The only randomised controlled trial to date in patients with severe depression was underpowered

What this study adds

This double blind randomised clinical trial showed that hypericum extract WS 5570 is at least as effective as paroxetine in ameliorating the symptoms of moderately or severely depressed patients

clinical trials of hypericum extract for different pharmaceutical companies. AD and MK are employees of Dr Willmar Schwabe Pharmaceuticals.

Ethical approval: The protocol was approved by the participating centres' appropriate independent ethics committees.

- 1 Linde K, Mulrow CD. St John's wort for depression. *Cochrane Database Syst Rev* 2004;(4):CD000448.
- 2 Harrer G, Hübner WD, Podzuweit H. Effectiveness and tolerance of the hypericum extract LI 160 compared to maprotiline: a multicenter double-blind study. *J Geriatric Psychiatry Neurol* 1994;7(suppl 1):S24-8.
- 3 Philipp M, Kohlen R, Hiller KO. Hypericum extract versus imipramine or placebo in patients with moderate depression: randomised multicentre study of treatment for eight weeks. *BMJ* 1999;319:1534-8.
- 4 Vorbach EU, Hübner WD, Arnoldt KH. Effectiveness and tolerance of the hypericum extract LI 160 in comparison with imipramine: randomized double-blind study with 135 outpatients. *J Geriatric Psychiatry Neurol* 1994;7(suppl 1):S19-23.
- 5 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.

- 6 Harrer G, Schmidt U, Kuhn U, Biller A. Äquivalenzvergleich Johanniskrautextrakt LoHyp-57 versus Fluoxetin. *Arzneimittel-Forschung* 1998;49:3-10.
- 7 Izzo AA. Drug interactions with St. John's Wort (*Hypericum perforatum*): a review of the clinical evidence. *Int J Clin Pharmacol Ther* 2004;42:139-48.
- 8 Montgomery SA. Clinically relevant effect sizes in depression. *Eur Neuropsychopharmacology* 1994;4:283-4.
- 9 Committee for Proprietary Medicinal Products. *Points to consider on switching between superiority and non-inferiority*. London: European Agency for the Evaluation of Medicinal Products, 2000.
- 10 Paykel ES. The classification of depression. *Br J Clin Pharmacol* 1983;15(suppl 2):155-9S.
- 11 Hypericum Depression Trial Study Group. Effect of *Hypericum perforatum* (St. John's wort) in major depressive disorder. *JAMA* 2002;287:1807-14.
- 12 Dunner DL, Dunbar GC. Optimal dose regimen for paroxetine. *J Clin Psychiatry* 1992;53(suppl):21-6.
- 13 Bourin M, Chue P, Guillon Y. Paroxetine: a review. *CNS Drug Rev* 2001;7:25-47.

(Accepted 17 December 2004)

doi 10.1136/bmj.38356.655266.82



Mortality data in adult cardiac surgery for named surgeons: retrospective examination of prospectively collected data on coronary artery surgery and aortic valve replacement

Ben Bridgewater on behalf of the adult cardiac surgeons of north west England

Abstract

Objectives To present named surgeon mortality for isolated first time coronary artery surgery and aortic valve surgery.

Design Retrospective analysis of prospectively collected data.

Setting All NHS hospitals undertaking adult cardiac surgery in north west England.

Participants 25 consultant surgeons carrying out coronary artery surgery and aortic valve replacement between April 2001 and March 2004.

Main outcome measures Mortality for both operations according to surgeon. EuroSCORE to stratify patients into low and high risk.

Results 10 163 patients underwent surgery under 25 surgeons. The average number of patients per surgeon was 363 for coronary artery surgery and 44 for aortic valve replacement. Seventeen per cent of the patients undergoing coronary artery surgery and half of those undergoing aortic valve surgery were considered high risk. The average mortality was 1.8% (range 0-3.8%) for coronary surgery and 1.9% (0-12.5%) for aortic valve surgery. Mortality for all surgeons fell below 99% control limits of the national mean for both operations.

Conclusions The presented mortality figures for the two cardiac operations fell within accepted limits for all surgeons. The division of outcomes according to low and high risk patients is imperfect but may help to inform the public about the complexities of this type of analysis and prevent surgeons avoiding high risk patients who may benefit from an operation.

Introduction

Recent years have seen a move towards increased openness and transparency in healthcare delivery. This has been accelerated by a series of events, including the Bristol Royal Infirmary inquiry into paediatric cardiac surgery deaths.¹ One recommendation of the inquiry was that patients must be able to see information about the relative performance of individual consultants operating within hospitals. The Society of Cardiothoracic Surgeons of Great Britain and Ireland therefore published a study in 2004 of activity and performance of all consultants undertaking adult cardiac surgery in the United Kingdom.² The history leading to this analysis and the underlying methods have been comprehensively described.³ The study was conducted on a single operation: first time isolated coronary artery surgery. Because of a lack of comprehensive data on which to perform a complete analysis that would allow adjustments to be made for differing case mix, the benchmarking was done on "crude" non-adjusted mortality data. The exact mortality for individual surgeons was not given, but instead surgeons were listed with a comment about whether they met the Society of Cardiothoracic Surgeons standards, which were defined as being acceptable if the surgeon fell within 99.99% confidence intervals of the national average.

Janet Smith has commented that the General Medical Council could be criticised for putting the interests of doctors before the interests of patients.⁴ When it comes to publishing mortality data for individual surgeons there is potentially a conflict between the interests of these two groups and the confidence intervals used recently by the Society of Cardio-

South Manchester
University Hospital,
Manchester
M23 9LT

Ben Bridgewater
consultant surgeon

Correspondence to:
B Bridgewater
ben.bridgewater@
smuht.nwest.nhs.uk

BMJ 2005;330:506-10