Clinical Study Synopsis

This Clinical Study Synopsis is provided for patients and healthcare professionals to increase the transparency of Bayer's clinical research. This document is not intended to replace the advice of a healthcare professional and should not be considered as a recommendation. Patients should always seek medical advice before making any decisions on their treatment. Healthcare Professionals should always refer to the specific labelling information approved for the patient's country or region. Data in this document or on the related website should not be considered as prescribing advice.
The study listed may include approved and non-approved formulations or treatment regimens. Data may differ from published or presented data and are a reflection of the limited information provided here. The results from a single trial need to be considered in the context of the totality of the available clinical research results for a drug. The results from a single study may not reflect the overall results for a drug.

Date of study report	27-Jul-2015 (PharmNet.Bund: STW3-VI-212-D-2008-III-Z CTR_Synopsis Final 1.1)
Study title	**Double blind, randomized, placebo-controlled, multicenter clinical trial to evaluate the relapse prevention of a hypericum extract in outdoor patients with moderate depressive episodes (major depression)**
Sponsor	Bayer [Study conducted by former Sponsor: Steigerwald Arzneimittelwerk GmbH, Germany]
Sponsor's study ID	20997
NCT number	tbd
EudraCT number	2008-001417-26
Study Phase	Phase 3
Indication	Moderate depressive episodes (major depression)
Study objectives	**Primary objective(s):** Study Phase 2: Superiority of STW3-VI to placebo by comparison of the relapse rates in both groups at the end of phase 2, measured by the total score of Hamilton Depression Scale (HAMD-17). **Secondary objective(s):** Study Phase 2: • Time between baseline and occurrence of first relapse event measured by: o Total score of HAMD-17 o Beck Depression Inventory (BDI) o Clinical Global Impressions (CGI) • Course of Disease • Safety and Tolerability of STW3-VI
Test drug	Hypericum extract STW 3-VI (Laif® 900, BAY98-7108)
Batchnumber(s)	C0901004, C0907009, C1003003, C1011006
Active ingredient(s)	Hypericum extract
Dose	1 x daily, 1 tablet, 900 mg (study Phase 1 und 2)
Route of administration	Oral

Duration of treatment	The total duration of study treatment per patient was 36 weeks. 12 weeks were open phase 1 and 24 weeks randomized and double-blind phase 2.
Reference drug **Batch number(s)** **[Dose]** **[Route of administration]** **[Duration of treatment]**	Placebo C0901004, C0907009, C1003003, C1011006 1 x daily, 1 tablet, 900 mg (study Phase 1 und 2) Oral The total duration of study treatment per patient was 36 weeks. 12 weeks were open phase 1 and 24 weeks randomized and double-blind phase 2.
Main inclusion criteria	Study Phase 1: • Age 18-70 years • Diagnosis according to DSM-IV-TR: 296.32: Outpatients with a recurrent major depressive disorder (presence of at least two or more Major depressive episodes), experiencing currently a moderate Major depressive episode and therefore requiring treatment • Criteria for Major depressive episode: o HAMD-17 total score of 20-24 Study Phase 2: • Phase 1 responders (= Phase 1 patients who have at least a 50% decrease in the HAMD-17 total score after 12 weeks of Laif® 900 therapy).
Study design	Double-blind, randomized, placebo-controlled, multicenter clinical trial
Methodology	The study design included 2 phases: The therapy responders of a first phase (acute therapy; verum only) were included in an immediately following second phase, which examined the preservation of the therapeutic success. The design, duration and primary outcome of this study were in line with the guideline EMEA CPMP / EWP / 518/97. Phase 1 spanned 12 weeks and was designed single-arm and open. During this period, a total of 4 visits at 4-week intervals were planned for the patients. The aim of this phase was to successfully treat the acute moderate-grade depressive episode with STW3-VI; Therapy responders from this phase were eligible for Phase 2. During the 24-week phase 2, which was performed in a two-arm, double-blind and randomized manner, a total of 7 visits took place every 4 weeks. The aim of this phase was to evaluate the relapse rate in the Phase 1 treatment responders.
Statistical methods	• Statistical error probability = 5%, two-sided (error of the first kind) • Evaluation of the target values for Full Analysis Set (FAS Collective) and Valid Case Analysis Set (VCAS Collective) • Univariate and multivariate analyzes

	• Examination of different strata (e.g. gender, age) The statistical evaluation for the target outcomes "HAMD-17" and "Global Assessment of Effectiveness" was carried out based on the full analysis set (FAS). Additional statistical evaluation was done for the valid case analysis set (VCAS), comparing the respective results. Intermediate analyzes were not performed in both phases. Phase 1: The evaluation of phase 1 was only exploratory in nature. Phase 2: All primary and secondary outcomes, as well as accompanying and safety variables were explored at first and after that evaluated descriptively. In particular, the structural similarity of the verum and placebo groups to the baseline of Phase 2 was examined. The primary endpoint (= relapse rate) was confirmed by non-parametric tests for unit values. Statistical tests were carried out nonparametric method. There were: • the Mann-Whitney U test comparing metric and ordinal variables between two study groups • the Wilcoxon test to examine changes in metric and ordinal variables within the overall collective and within each study group • the Chi² test is used to compare shares. The comparison between more than two (sub-) groups was done with the Friedman or Kruskal-Wallis test. In the case of small sample sizes or imbalances between groups, exact tests were used for these tests. This was especially true for subgroup analyzes. Furthermore, survival times achieved (time until the onset of the first relapse) were graphically displayed in the form of Kaplan-Meier curves and statistically compared by means of the Log-Rank test and the Breslow test. The analysis of the follow-up was carried out in advance. In the evaluation, center effects, gender and age-specific influences and the effects of the HAMD value on baseline in phase 2 were investigated for the primary endpoint value. If such influences could not be excluded, the test results were supplemented by subgroup comparisons (centers, gender, age) for an exploratory analysis. All p-values from statistical tests in the area of exploratory analysis, which go beyond the examination of the primary endpoint, were also exploratory only.
Early termination	Not applicable
Substantial protocol changes	The final study protocol from 28.03.2011, includes 2 substantial amendements: *Protocol -amendment 1 from date* 22.06.2009 *introduced the following changes*: 1. Increase the number of trial centers: Originally 6 study centers with 6 investigators were planned. Due to the slow patient recruitment, which significantly increases the clinical part of the clinical trial, the number of centers was increased to 11 with 31 investigators, of which 2 study centers (day clinics) with 8 investigators could not include

	any patients. 2. Modification of an exclusion criterion Originally, any kind of psychotherapy was not allowed. As in this case however, supportive / supportive-adaptive psychotherapy procedures were not possible either, the regulation was adapted to: "revealing (depth psychology-based) psychotherapy procedures and cognitive behavioral procedures are not allowed." 3. Specification of the relapse definition Originally: Relapse = increase of the HAMD-17 total score by> 5 points towards baseline of Phase 2. After the first patients had finished Phase 1 so low in the HAMD-17 aggregate score, it became obvious, that with an increase of 6 points, they would have been still in the depression-free area. Therefore, the relapse definition was changed to: Relapse = Increase of HAMD-17 total scores are> 5 points higher than baseline in Phase 2, provided that a value of ≥ 12 points (range of mild depression) was reached during the phase 2. *Protocol -amendment 2 from date 28.03.2011 introduced the following changes*: 4. Increase in the number of cases for Phase 2 The study plan originally involved a case number of 99 patients per group in Phase 2 which a presumed a drop-out rate of 10%. As the number of drop-outs without relation to relapse, that could not be used for the evaluation of the relapse rate, was higher in the ongoing clinical trial than assumed and furthermore no valid forecast for these cases was possible, the initially assumed drop-out rate of 10% was increased to 25%, increasing the required number of cases from 99 to 120 patients per group.
Study period	Study Start Date: 03-Mar-2009 Study End Date: 22-Nov-2011
Study center(s)	The study was conducted at 9 study sites in Germany.
Number of subjects	Planned: 400 Analyzed: 398

Criteria for evaluation	
Efficacy	Primary outcomes: • Relapse rate: Definition of the Relapse Event: Increase of the HAMD-17 total score in Phase 2 by > 5 points compared to Phase 2 baseline, if at the same time a total score of at least 12 is reached. Secondary outcomes: • Time until the first relapse event occurs • HAMD-17 total score history • Beck Depression Inventory (BDI) • Clinical Global Impression (CGI) • Global assessment of efficacy by investigators and patients at the end of each study phase
Safety	Safety and Tolerability: • Occurrence of adverse events • Global assessment of tolerability by investigators and patients at the end of each study phase • Clinical and clinical-chemical parameters (heart rate, blood pressure, body weight, laboratory parameters of blood and urine)
Clinical pharmacology	Not applicable

Subject disposition and baseline

The present multicenter clinical trial was performed in 9 study centers with 23 active investigators. In the period from 03-Mar-2009 to 03-May-2011 398 patients (260 = 65.3% women and 138 = 34.7% men) were included in the clinical trial (phase 1 / acute therapy), taking into account the inclusion and exclusion criteria. The patients reported at least one depressive episode in their psychiatric history (mean 3.9 +/- 5.4 episodes, maximum 62 episodes). Of the 398 patients, 388 (= 97.5%) entered the safety group, 375 (= 94.2%) the FAS and 330 (= 82.9%) the VCAS collective. 395 (= 99.2%) of the 398 patients indicated that they were "ethnically" Caucasian. 240 Phase 1 responders were accepted into the double-blind, randomized, placebo-controlled Phase 2 trial. Of these, 237 (98.8%) were in the safety group, 234 (97.5%) in the FAS group and 207 (86.2%) in the VCAS group. All patients in study phase 2 were Caucasian.

Efficacy

Phase 1

On average, the total HAMD-17 score decreased by 62.4% in the FAS and 65.7% in the VCAS collective.

While comparing visit 4 with visit 1, 79.9% (298 out of 373) of the patients in the FAS collective and 85.4% (281 out of 329) of the patients in the VCAS collective showed a decrease in the HAMD-17 total score of at least 50% and were therefore by definition responders.

At the end of Phase 1, the BDI showed a mean decrease of 55.7% in the FAS collective.
At the end of Phase 1, according to CGI severity, 68.1% (243 out of 357) of the patients were diagnosed as "not" or "borderline" diseased (FAS collective).
According to CGI status changes, 89.9% (321 out of 357) of patients in the FAS collective did better, most of them "much better" (58.8%).
According to the CGI Efficacy Index, the therapeutic efficacy was rated "very good" in 59.4% (212 out of 357) of the patients in the FAS collective; 94.4% (337 out of 357) of the patients had no therapeutic risks.
Globally, in the FAS collective, the investigators and patients, respectively, rated the effect 81.5% and 77.0% "very good" or "good" at the end of Phase 1, respectively.

Primary variable:

Phase 2:

Relapsrate:
During Phase 2, 66 relapse events occurred in the FAS collective (28.2%, 66 out of 234), of which 27.1% (32 of 118) were in the verum group and 29.3% (34 of 116) in the placebo group (p = 0.772). Of the total 60 (29.0%, 60 out of 207) events in the VCAS collective, 29 were in the verum group and 31 in the placebo group (p = 1.000).
Noteworthy is the much greater relapse frequency in women compared to men (34.4 vs. 15.6%, FAS collective). On the other hand, an age dependency did not seem to exist.

Secondary variables:

Time until the 1st relapse event
In the verum group (STW3-VI), the 1st relapse occurred on average 16.6 days (FAS) or 16.4 days (VCAS) later than in the placebo group (p = 0.192 in the FAS collective, p = 0.239 in the VCAS collective) and was less severe on the HAMD-17 scale.
No statistically significant group differences were recorded in the HAMD-17 and BDI total score as well as in the CGI at any time of the visit (FAS collective).
Globally, in the FAS collective, the investigators and patients, respectively, rated efficacy in the verum group at 76.5% and 74.8% "very good" or "good" in the placebo group at the end of Phase 2, respectively to 78.2% and 70.9%, respectively.

The following table summarizes important Phase 2 efficacy data:

Parameter	STW3-VI	Placebo	Difference	p-value
Relapse-Rate (%): FAS	27,1	29.3	2.2	0.772
Time to 1st Relapse (mean, days): FAS	94.6	78.0	16.6	0.192
HAMD-17 at 1st Relapse (mean, points): FAS	16.4	17.6	1.2	0.160
HAMD-17 min. 20 points at 1st Relapse: FAS	18.8	38.2	19.4	0.233
Patients with discontinued therapy after Relapse (total, %): FAS	34.4	67.6	33.2	0.013

Safety

Adverse events (AE)

In total, 304 adverse events were registered in 219 patients. This included 183 AEs in 139 patients in Phase 1, including one AE rated severe (acute sudden hearing loss, with no causal relationship with the study medication). Phase 2 involved 121 adverse events in 80 patients: 66 events in the verum group (42 patients) and 55 events in the placebo group (38 patients), including 5 severe AEs (3 on STW3-VI, 2 on placebo) each without any connection with the intake of the study medication. There was no statistically significant difference between the treatment groups in phase 2 for the proportion of patients with an AE.

Adverse drug reaction (ADR)

Of the 304 AEs, a total of 95 in 61 patients were declared as ADRs by STW3-VI. Of these, in Phase 1 there were 64 ADRs in 48 patients. In total, there were 12 drop-outs associated with ADRs. Phase 2 included 31 ADRs in 13 patients taking STW3-VI; one drop-out was related to a ADR. There were no serious ADRs. The ADRs are already known predominantly as side effects of the skin and gastrointestinal tract. In the placebo group, ADRs were reported in four patients.
Globally, in the FAS collective, the investigators / patients rated the tolerability in phase 1 as 94.8% and 93.6% as "very good" or "good" and in phase 2 in the verum group as 97.4%. 94.8% were "very good" or "good", 99.1% and 96.3% respectively in the placebo group.

Safety parameters

- *Clinical parameters:* body weight, blood pressure, heart rate.
- *Clinical-chemical parameters:* small blood count, enzymes ALAT, ASAT, γ-GT, metabolites bilirubin, creatinine, uric acid and glucose and protein in the urine.

All parameters changed little on average, as well as during Phase 1 and Phase 2.

Conclusions from the results of Phase 1 and 2:

Moderately depressed patients benefited from the 12-week acute STW3-VI treatment, as a mean reduction of the HAMD-17 score of 62.4% (FAS) and a response rate of 79.9% (FAS) was achieved. This basically confirms efficacy data from previous studies with STW3-VI. However, in contrast to expectation and planning and due to its inherent methodological difficulty regarding the protocol

definition of a relapse event, this study showed no significant reduction in relapse frequency among STW3-VI versus placebo.
From the secondary outcome criteria, the following trend could be deduced:

- The relapse recorded with therapy with STW3-VI occurred later compared to placebo and was less severe.
- Furthermore, STW3-VI is also well tolerated in an application over 36 weeks.

Reanalyzes of the unexpected result of the study
Reanalyzes of the unexpected result of the study was likely due to methodological difficulties - The following methodological aspects were considered:
The primary study endpoint was susceptible to the definition of a relapse event. The threshold of 12 points as selected here was too low, and also against the background definition of the German National Care Guideline for Unipolar Depression, considering the middle-grade depressive syndrome in a HAMD 17 range of 17-24 points. It would thus be plausible to speak of a relapse of the disease in moderately depressed patients, provided the severity of this disease, previously found in the same episode, was at least approximately reached again (inclusion criterion HAMD-17 of 20-24).
However, a clear regulatory definition of the relapse event is missing, and there is also a variety of information in the scientific literature on how to capture a relapse event. For example, some authors suggest a combination of HAMD increase and individual psychiatric assessment of patients.
Exploratory reanalyzes of this study indeed indicated that a better definition of the relapse criterion would be able to demonstrate a significant superiority of STW3-VI over placebo. Even a simple increase in the HAMD-17 threshold for relapse resulted in an increasing numerical superiority of STW3-VI over placebo. For the present study, no confirmatory results can be derived from this. However, this reanalysis underlines that patients treated with placebo have an increased risk of developing more severe and therefore clinically relevant relapses, see the table below:

Threshold	**Relapse Verum (V) Number (%)**	**Relapse Placebo (P) Number (%)**	**Ratio P : V**
12	32 (27,1%)	34 (29,3%)	1,062
13	29 (23,7%)	33 (28,4%)	1,138
14	27 (22,9%)	31 (26,7%)	1,148
15	24 (20,3%)	31 (26,7%)	1,292
16	21 (17,8%)	28 (24,1%)	1,333
17	19 (16,1%)	26 (19,8%)	1,368
18	12 (10,2%)	23 (19,8%)	1,917
19	11 (9,3%)	17 (14,7%)	1,545
20	8 (6,8%)	16 (13,8%)	2,000
21	4 (3,4%)	10 (8,6%)	2,500

22	3 (2,5%)	9 (7,8%)	3,000
23	2 (1,7%)	5 (4,3%)	2,500
24	1 (0,8%)	3 (2,6%)	3,000

An exploratory post-hoc relapse definition examined a combination of HAMD-17 increase and medical-psychiatric assessment as mentioned above.
The patient records of patients with relapses according to protocol definition (HAMD criteria) were examined by a psychiatrist for events that pointed to interference factors due to medical or environmental influences according to the classification criteria of the DSM-IV-TR, making a real relapse of depression unlikely. In this analysis, there was a relapse incidence of 26.7% in the placebo group and 14.4% in the STW3-VI group ($p = 0.023$, not final trial report confirmatory). This suggests a definition of the relapse for follow-up studies, which provides for a medical assessment of the patient by the examiner, taking into account the axes and classification criteria of the DSM, and possibly the evaluation by an independent end-point committee.

Overall conclusions

Overall, the study showed a low relapse rate not expected in the planning in both treatment groups, with the very low relapse rate of 29.3% in the placebo group being explained by a high placebo effect. This low value was well below the 41% in the meta-analysis by Geddes et al. 2003, which was used for the sample size planning. One possible explanation for this large difference is that this meta-analysis included studies of very different duration (0.5-3 years), and the placebo relapse rate increased with increasing study duration. For example, the two most recent studies in this meta-analysis, also with a placebo-controlled treatment period of about half a year, found a relapse incidence of placebo of only 26.7%, comparable to the relapse frequency in this study (29.3%) and the same in another Hypericum study (25.7%).
These data were not taken into account in the case number estimation for the present study or had not yet been published. In order to demonstrate a significant superiority of STW3-VI over placebo in the withdrawal-design of this study, a higher number of patients or a longer relapse-prevention phase than 24 weeks would have been required.

Publication(s) based on the study

None at the time of this report.