PHYTOTHERAPY RESEARCH
Phytother. Res. **15**, 344–350 (2001)
DOI: 10.1002/ptr.981

Efficacy and Tolerability of a Standardized Willow Bark Extract in Patients with Osteoarthritis: Randomized Placebo-controlled, Double Blind Clinical Trial†

B Schmid,[1] R Lüdtke,[2] H-K Selbmann,[2] I Kötter,[3] B Tschirdewahn,[4] W Schaffner[5] and L Heide[1]*

[1]Pharmazeutisches Institut, Universität Tübingen, Germany
[2]Institut für Medizinische Informationsverarbeitung, Universität Tübingen, Germany
[3]Medizinische Universitätsklinik, Tübingen, Germany
[4]Federseeklinik, Bad Buchau, Germany
[5]Department Pharmazie, Universität Basel, Switzerland

This study assessed the clinical efficacy of a chemically standardized willow bark extract in the treatment of osteoarthritis. Willow bark extract, in a dose corresponding to 240 mg salicin/day, was compared with placebo in a 2-week, double-blind, randomized controlled trial. The primary outcome measure was the pain dimension of the WOMAC Osteoarthritis Index. Secondary outcome measures included the stiffness and physical function dimensions of the WOMAC, daily visual analogue scales (VAS) on pain and physical function, and final overall assessments by both patients and investigators. A total of 78 patients (39 willow bark extract, 39 placebo) participated in the trial. A statistically significant difference between the active treatment and the placebo group was observed in the WOMAC pain dimension ($d = 6.5$ mm, 95% C.I. = 0.2–12.7 mm, $p = 0.047$); the WOMAC pain score was reduced by 14% from the baseline level after 2 weeks of active treatment, compared with an increase of 2% in the placebo group. The patient diary VAS confirmed this result, and likewise the final overall assessments showed superiority of the willow bark extract over the placebo (patients' assessment, $p = 0.0002$; investigators' assessment, $p = 0.0073$). It is concluded that the willow bark extract showed a moderate analgesic effect in osteoarthritis and appeared to be well tolerated.

Keywords: osteoarthritis; willow bark; *Salix*; herbal remedies; WOMAC.

INTRODUCTION

Osteoarthritis (OA) is the most common form of arthritis, and is also probably the most common disease affecting humans. Radiological studies suggest that more than 80% of people over 65 have OA (Mankin, 1989). Drug therapy of OA is empiric and largely directed towards providing symptomatic relief, primarily by the use of analgesics and non-steroidal antiinflammatory drugs (NSAIDs). Over the past years, herbal medicines have become increasingly popular, and there is a growing number of OA patients who are unwilling to comply with long-term NSAID treatment and who wish to use herbal antirheumatic medicines. Although there is a number of such remedies on the market, few have been chemically standardized in their composition, and there is a definite lack of published evidence of their efficacy as proven by randomized controlled trials.

* Correspondence to: Professor L. Heide, Pharmazeutisches Institut, Universität Tübingen Auf der Morgenstelle 8, 72076 Tübingen, Germany.
E-mail: heide@uni-tuebingen.de
† First published in German language in *Zeitschrift für Rheumatologie* 2000; **59**: 314-320.
Contract/grant sponsor: Alfried Krupp von Bohlen und Halbach Foundation.
Contract/grant sponsor: Karl and Veronica Carstens Foundation.

In Germany, a systematic evaluation of herbal medicines has been carried out by the Commission E of the Federal Health Agency (now called the Federal Institute for Drugs and Medical Devices). Commission E has prepared monographs on selected herbal medicines, defining their therapeutic uses and dose, and has set standards for their content of active principles when possible (Blumenthal, 1998). The only antirheumatic herbal drug for internal use which has been approved by the Commission E, with a chemical standardization for an active constituent, is willow bark.

Willow bark is commonly regarded as one of the first examples of modern drug development from a herbal remedy: 160 years ago, the chemical oxidation of salicin, a known constituent of the bark of the willow tree (Latin name: Salix), resulted in a new substance termed 'salicylic acid', and the acetylated derivative of this substance eventually turned into one of the most successful drugs in history, Aspirin™. The starting point for this development was the antiinflammatory, antipyretic and analgesic effects of willow bark, mentioned already in the classical antiquity (Aronson, 1994) and reported in detail for the first time by Stone in 1763. Today, willow bark is included in the German pharmacopoea, and a monograph for the European pharmacopoea is in preparation. Commission E of the German

Received 1 November 2000
Accepted 1 November 2000

Federal Health Agency has approved the use of willow bark preparations for the treatment of 'diseases accompanied by fever, rheumatic ailments and headaches' (Blumenthal, 1998), in a daily dose equivalent to 60–120 mg salicin. The European Scientific Cooperative on Phytotherapy (ESCOP) recommends an equivalent of up to 240 mg salicin/day (ESCOP, 1997). Besides two anecdotal reports (Mayer and Mayer, 1949; Weyers, 1992) only one randomized controlled trial on the efficacy of a non-standardized willow bark preparation has been published in detail (Mills *et al.*, 1996).

Our interest was prompted by a small ($n = 20$) pilot study (Schaffner, 1997) which indicated a clinically relevant analgesic effect of willow bark extract in a dose equivalent to 240 mg salicin/day. Salicin and related constituents (Meier *et al.*, 1985; Julkunen-Tiitto, 1986) of willow bark are, after oral application, metabolized to salicylic acid (Steinegger and Hövel, 1972; Pentz *et al.*, 1989). However, even assuming 100% bioavailability, 240 mg salicin would produce no more than 115 mg salicylic acid, and would therefore not be expected to produce analgesic or antiinflammatory blood levels of salicylic acid. If the reports on the clinical efficacy of willow bark are true, other constituents of the plant may contribute to the overall effect. In this case, willow bark extract, or active constituents thereof, may potentially offer an analgesic therapy with a better tolerability (e.g. less gastrointestinal side effects) than acetylsalicylic acid or NSAIDs.

We have therefore carried out a randomized placebo-controlled clinical trial to investigate the analgesic effect of a chemically standardized willow bark extract in patients with osteoarthritis. We have applied the validated WOMAC Osteoarthritis Index (Bellamy, 1988) to assess drug efficacy, and followed the most recent recommendations for outcome measures in phase III clinical trials in OA (Bellamy *et al.*, 1997).

Methods

Study medication and blinding. At the beginning of the study, no commercial preparation was available containing only willow bark and corresponding to the dose given by the Commission E of the German Federal Health Agency. Therefore, the study medication was produced by the authors. An extract of the bark of *Salix purpurea x daphnoides* was kindly provided by Salushaus GmbH (Bruckmühl, Germany). Chemical analysis (Meier *et al.*, 1985) showed 17.6% total salicin. Coated tablets containing 340 mg extract (corresponding to 60 mg salicin) were produced with the help of Zeller AG (Romanshorn, Switzerland). Coated placebo tablets were produced in the same way, containing cellulose and lactose. For both medications, the coating consisted of Opadry yellow and Macrogol 20,000. The colour, size, odour and taste of all tablets was identical.

Randomization and study design. The protocol was approved by the ethical committee of the Tübingen University Hospital and carried out in accordance with the GCP guidelines. Participants were recruited from in-patients of a rheumatology hospital (Federseeklinik, Bad Buchau, Germany). The following inclusion and exclusion criteria were applied: Inclusion—osteoarthritis of the hip or of the knee, verified according to the clinical, laboratory and radiographic criteria of the ACR (Altman, 1991) age >18 years (men) or >50 years (women); and written informed consent; Exclusion—Known allergic reactions to salicylates; kidney failure (blood creatinine >84 μmol/L); abnormal liver function (SGOT >35 U/L or SGPT >35 U/L or GGT >50 U/L); gastrointestinal ulcers (bleeding or discoloured stool during the past 8 weeks); malignant diseases; systemic therapy with corticosteroids during the past 8 weeks; surgery of the test joint during the past 8 weeks; inflammatory joint diseases (ESR >40 mm/h); chronic heart failure (NYHA grade III or IV); chronic obstructive airway disease requiring prophylactic medication; participation on a clinical trial during the past 4 weeks; lack of compliance with placebo medication or patient diary in the wash-out phase.

Participants were grouped in four strata according to the location of the osteoarthritis (hip or knee as the most severely affected joint) and according to the two involved departments of the hospital. Randomization occurred in blocks of four within each stratum, using computer generated random numbers (Excel 5.0). Both assessors and patients were blind to the allocation and not informed about the block size until after completion of the study. After a wash-out phase of 4–6 days with placebo (day −4 to day 0), participants were treated for 2 weeks (days 1–14) with either willow bark extract or placebo (two tablets twice daily; willow bark extract corresponding to a dose of 240 mg salicin/day). Study medications were taken in the morning and at noon, half an hour before meal times. No additional analgesics, NSAIDs or systemic corticosteroids were allowed during the wash-out and study phases. Patients were assessed by the medical doctor at days −4, 0, 7 and 14. The initial assessment (day −4) comprised medical history and examination, Lequesne Index (Lequesne, 1991) and WOMAC-VA 3.0 Osteoarthritis Index (Bellamy *et al.*, 1988). We followed the translation by Stucki *et al.*, (1996) for the German phrasing of the questionnaire. At the first visit, the most painful hip or knee was identified as the key joint (Bellamy *et al.*, 1990) i.e. this joint was the primary focus of measurement for future assessment in the study. Blood and urine samples were taken for standard laboratory tests, and all patients were instructed to take placebo tablets during the wash-out phase. Patients kept a diary during the entire study in order to document the number of tablets taken as well as any adverse events, and in order to rate every evening both pain and physical function on 100 mm visual analogue scales.

At the second study visit (day 0), laboratory findings were compared with exclusion criteria, compliance to placebo medication was verified by tablet counting, and diary entries were checked. Patients who met all study criteria filled in the WOMAC questionnaire (baseline), received study medication and thereby entered the intent-to-treat population.

At the study visits on days 7 and 14, patients filled in WOMAC questionnaires, and compliance was checked by diary entries and tablet count. Adverse events were recorded by checking diary entries as well as by direct questioning of the patients. Additional WOMAC questionnaires were filled in by the patients independently on day 3 and day 10. At the study visit on day 14, blood and urine samples were taken, and patient and physician

Table 1. Comparison of characteristics at enrolment

Characteristic	Willow bark extract ($n = 39$)	Placebo ($n = 39$)
Age (years)	52.4 ± 7.0	53.5 ± 10.5
Sex (M/F)	30/9	29/10
Height (cm)	170.7 ± 7.0	171.2 ± 10.2
Body weight (kg)	80.4 ± 10.4	85.6 ± 18.0
OA of the hip/knee	22/17	22/17
OA bilateral/unilateral	25/14	29/10
Department of hospital (A/B)	36/3	35/4
Duration of disease (years)	9.3 ± 7.4	7.3 ± 4.6
Lequesne score (day −4)	8.9 ± 3.0	9.6 ± 3.8
WOMAC pain score (day 0)	34.1 ± 19.3	44.1 ± 26.5
Prior use of analgesics	10 yes/ 29 no	18 yes/20 no*[a]

[a]Values are mean ± SD.
*One value missing.

Table 2. Physical therapy during study phase

Type of physical therapy	Number of sessions: Willow bark extract group	Number of sessions: Placebo group
Exercise therapy	17.6	17.2
Mud bath therapy	3.7	3.7
Massage therapy	3.6	3.6
Electrotherapy	1.4	0.9

independently recorded their final overall assessment of the change of disease activity by the study medication on 100 mm visual analogue scales.

All in-patients of the study hospital received regular physical therapy (Table 2). Each physical therapy session followed standard procedures of the study clinic.

Statistical analysis. The primary outcome variable was the difference in the WOMAC pain dimension between day 0 and day 14. Differences in the WOMAC stiffness and physical function dimensions, daily VAS scores on pain and physical function, and the patients' and doctors' final overall assessment of the effect of study medication on disease activity were used as secondary endpoints (Bellamy *et al.*, 1997; Hochberg *et al.*, 1997). Throughout this paper, scores for all three WOMAC dimensions are uniformly reported as mm on a 0–100 mm scale. The statistical method used for the analysis of the primary and secondary outcome variables was the *F*-test in a two-factorial ANOVA (analysis of variance) using the criteria defining the four strata (see above) as block factors. In the case of missing values the last available value was used. Data were analysed both on an intent-to-treat basis, including all 78 patients entering the study phase, and on a per-protocol basis, excluding 10 patients who committed protocol violations (see Results). Patients who withdrew during the study phase were included in the analysis with their WOMAC Index taken at the day of withdrawal, before any additional analgesic therapy was started. Patients who withdrew due to pain and the requirement for additional analgesic therapy, or withdrew after at least 7 days of study medication, were included in both intent-to-treat and per-protocol data analysis, as defined in the protocol.

In all statistical tests, the level of Type I error (2-tailed) was set at 0.05. Reported *p*-values for secondary outcome variables and adverse events are considered as descriptive only.

Results

Some 165 patients with a medical history of osteoarthritis of the hip or/and the knee were screened, and 86 were enrolled and entered the wash-out phase. The investigators withdrew six patients before they entered the study phase, since their laboratory values fulfilled the exclusion criteria of the protocol. Two patients withdrew due to adverse events (one gastrointestinal pain, one vomiting and migraine). A total of 78 were finally randomized to take either placebo or willow bark extract and formed the intent-to-treat population (Table 1). During the study phase, four patients (three placebo, one active treatment) withdrew due to pain and a requirement for additional analgesic therapy, and one patient (active treatment) withdrew on day 14 due to allergic symptoms. All five withdrawals were included in the data analysis with their WOMAC score taken on the day of withdrawal (see Methods).

Important protocol violations were committed by ten patients: eight (three placebo, five active treatment) received electrotherapy at the key joint after day 7, one placebo patient had the baseline WOMAC Index taken on day 2 (instead of day 0), and one active treatment patient underwent only a 1 day wash-out phase (instead of 4 days). These ten patients were excluded from the per-protocol analysis.

Efficacy

The validated WOMAC-VA 3.0 questionnaire contains a total of 24 visual analogue scales (100 mm each), five of which refer to the pain dimension, two to stiffness and 17 to physical function (Bellamy *et al.*, 1998). During the placebo wash-out phase, all three WOMAC dimensions showed an improvement (Table 3). During the following study phase, only the active treatment group showed a further improvement of pain and physical function, the clinically most important dimensions of the WOMAC Index (Bellamy *et al.*, 1988; 1997). The visual analogue scales on pain and physical function in the patient diaries confirmed these results. A statistical analysis of the data in Table 3 was carried out for the validated WOMAC outcome measures, using an ANOVA as defined in the protocol; ANOVA estimates are therefore not identical to the difference of means in Table 3. A statistically significant superiority of active treatment over placebo was observed for the WOMAC pain score, i.e. the primary outcome measure (intent-to-treat analysis: difference 6.5 mm, 95% C.I. = 0.2–12.7 mm; $p = 0.047$). As expected, the analysis of the WOMAC pain score on a per-protocol basis showed the difference between the active treatment and placebo more clearly than the intent-to-treat analysis (Fig. 1); the ANOVA estimate of the difference between the groups is 7.3 mm (95% C.I. = 1.3–13.4 mm; $p = 0.0196$). Physical function improved under active treatment, but the difference to the placebo group

Table 3. Mean values for efficacy outcome measures

Variable	Group	Baseline ± SD					Termination	Within group baseline vs termination difference of means	Between groups baseline vs termination difference estimate by ANOVA
WOMAC Index		Day −4	Day 0	Day 3	Day 7	Day 10	Day 14		
Pain	P	48.3	44.1 ± 26.5	43.0	43.7	43.5	45.1	+ 1.0 (+2 %)	6.5
	W	37.4	34.1 ± 19.3	31.0	30.4	29.5	29.3	− 4.8 (−14 %)	(p = 0.047)
Physical function	P	46.9	43.8 ± 24.2	42.4	43.3	42.8	44.8	+ 1.0 (+2 %)	3.8
	W	41.0	37.3 ± 22.2	35.6	35.0	32.7	33.8	− 3.5 (−9 %)	(p = 0.112)
Stiffness	P	47.1	44.9 ± 26.2	42.0	41.9	40.3	41.2	− 3.7 (−8 %)	1.3
	W	43.7	37.0 ± 25.8	35.2	34.1	32.9	33.2	− 3.8 (−10 %)	(p = 0.715)
Patient diary VAS		Day −3	Day 0	Day 3	Day 7	Day 10	Day 14		
Pain	P	51.5	48.2 ± 27.7	44.4	46.2	46.8	48.2	± 0 (±0 %)	n.d.[a]
	W	50.5	43.2 ± 27.4	39.7	37.5	34.2	33.2	−10.0 (−23 %)	
Physical function	P	42.6	41.5 ± 29.2	41.4	40.8	38.6	41.3	− 0.2 (−0.4 %)	n.d.[a]
	W	44.5	43.8 ± 30.1	40.3	36.4	36.6	34.2	− 9.6 (−22 %)	

Scores for all three WOMAC dimensions and for patient diary VAS are uniformly reported as mm on a 0–100 mm scale. Data refer to the intent-to-treat analysis. P, placebo; W, willow bark extract; n.d., not determined.

[a]A statistical analysis of the baseline vs termination difference was only carried out for the validated WOMAC Index, as defined in the study protocol.

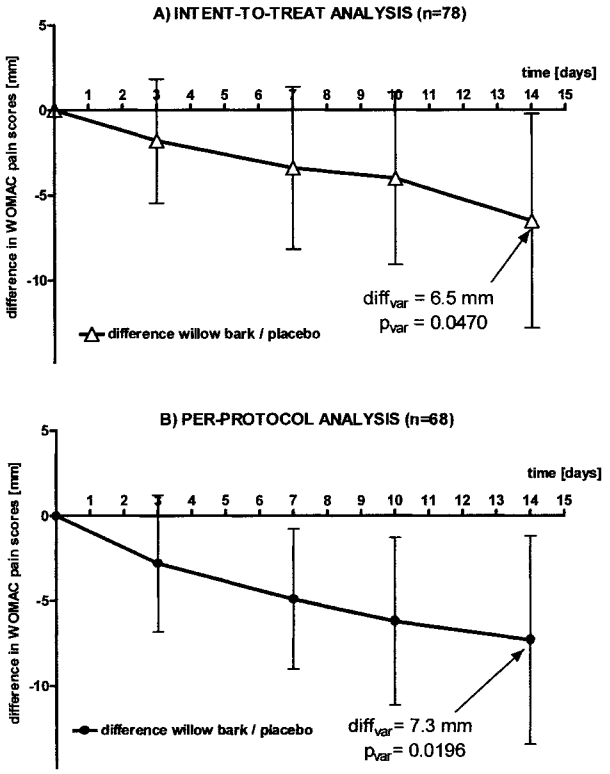

Figure 1. Improvement of WOMAC pain scores versus baseline: estimated difference between willow bark extract and placebo group by ANOVA (analysis of variance) with 95% C.I. Scores are reported on a 0–100 mm scale.

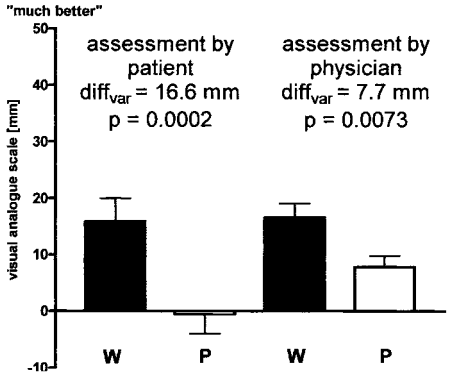

Figure 2. Final overall assessment of the change of disease activity by study medication. 100 mm visual analogue scale (+50 mm = 'much better', −50 mm = 'much worse'). Mean values and 95% C.I. are shown. P, placebo; W, willow bark extract.

was not statistically significant. Stiffness showed a marked improvement in the active treatment group during the wash-out phase; during the following study phase, improvements in both groups were nearly identical. MANOVA of the outcomes observed in the four strata of the study showed no relevant differences between the hip and knee OA, nor between the two departments of the hospital.

Despite the randomization, the placebo group showed a higher baseline WOMAC score than the active treatment group, as well as a higher proportion of patients who had used analgesics prior to the study (Tables 1 and 3). This difference was largely due to four patients in the placebo group who showed baseline WOMAC pain scores of >90 mm, while all other patients in the study had baseline scores of <80 mm. If both the initial WOMAC pain score (day 0) and the prior use of analgesics were included as covariates in an ANCOVA (analysis of covariance), the difference of the primary outcome measure between the active treatment group and the placebo group resulted in the intent-to-treat population as 7.5 mm (95% C.I. = 0.9–14.0 mm; $p = 0.025$), and in the per-protocol population as 8.8 mm (95% C.I. = 2.7–14.8 mm; $p = 0.005$). Therefore, the initial differences between both treatment groups probably led to a slight underestimation of the outcome difference between the two groups in the original ANOVA defined in the protocol.

The final overall assessment of the effect of the study medication on the disease activity (Fig. 2) clearly showed superiority of the active treatment over the placebo. Improvement under active treatment was independently rated by both patients and physicians as approximately +17 mm on a visual analogue scale, with +50 mm being the maximum possible improvement. The patients reported no change of disease activity by the placebo, and the physicians also reported a much lower improvement under the placebo than under the active treatment. The between-group differences were highly significant (patients' assessment, $p = 0.0002$; physicians' assessment, $p = 0.0073$).

At the final visit, each patient was asked which type of medication (placebo or willow bark extract) he believed he had received in the preceeding days. In the placebo group, 12 patients believed they had taken the willow bark extract, 10 believed they had taken the placebo, and 17 did not have a supposition about which study group they belonged to. In the active medication group, 14 patients believed they had indeed taken the willow bark extract, six believed they had taken the placebo, and 19 had no supposition. The fact that in the willow bark group patients supposed somewhat more frequently that they had taken the active treatment than did the placebo group would be expected for an effective study medication. In total, these data indicate successful blinding during the study.

Tolerability and safety

The frequency and nature of adverse events reported during the study are illustrated in Table 4. In both treatment groups the most frequent clinical adverse events reported were allergic skin reactions and gastrointestinal disorders. As shown in Table 4, 16 patients (41%) of each treatment group reported one or more adverse events. Numerically, there were more adverse events reported in the placebo group ($n = 28$) than in the active treatment group ($n = 17$). The most important adverse event in the active treatment group was a skin rash starting in a patient on day 10 and causing withdrawal from the study on day 14; this patient had a medical history of frequent allergic reactions, however, not to salicylates.

Changes in haematology, clinical chemistry and urinalysis data (day −4 versus termination) showed statistically significant between-group differences for white blood cell count (active treatment: $-\ 0.6 \times 10^3/\mu L$; placebo: $-0.01 \times 10^3/\mu L$), SGOT (active treatment: −0.26 U/L;

Table 4. Adverse events by the two treatment groups

Event	Willow bark extract ($n = 39$)	Placebo ($n = 39$)
Skin and appendages disorders	6	5
GI system disorders	3	7
Infections	2	2
Headache/migraine	1	2
Change in haemogram	1	2
Hypertriglyceridaemia	1	1
Musculoskeletal pain	0	2
Sleeplessness	0	2
Other adverse events	3	5
Total number of adverse events (AE)[a]	17	28
Total number of patients experiencing AE	16	16

[a] Some patients had more than one AE.

placebo: −0.94 U/L) and glucose in serum (active treatment: −3.76 mg/dL; placebo: −11.03 mg/dL). The difference in serum glucose could be explained by two patients of the placebo group starting antidiabetic treatment during the study. None of these changes was considered clinically relevant, and all mean values of either treatment group were within normal ranges throughout the study.

Within the haematology, biochemistry and urinanalysis parameters examined, 15 placebo and 15 active treatment patients showed one or more laboratory values out of the normal range at termination, but the majority of these patients had shown abnormal values already at entry. Normal starting and significantly abnormal termination values were reported for only four patients: one active treatment and two placebo patients showed changes in haemogram, and one of these placebo patients as well as another active treatment patient showed hypertriglyceridaemia.

Discussion

This study shows that the willow bark extract, in the composition and dose used in our trial, was efficacious as an analgesic in the treatment of OA of the knee and the hip. This conclusion is supported by the majority of the secondary outcome measures, most prominently by the patients' overall assessment (Fig. 2). The results of the different outcome measures of this study were highly consistent.

We used a high dosage of willow bark extract, corresponding to the upper limit of the recommendation of the European Scientific Cooperative on Phytotherapy (ESCOP). This dosage can not be achieved with several of the commercial willow bark preparations on the market, and therefore the results of our study will only apply to preparations of comparable composition. Independently from our study, three different willow bark preparations containing 60 mg salicin per tablet (like the study medication in this report) have recently become available on the German market, e.g. Assplant™ (Robugen GmbH, Esslingen/Germany), which is produced from the same willow bark extract as the medication prepared for the present study.

Parallel to the present clinical trial, we have carried out a pharmacokinetic trial with ten healthy volunteers receiving willow bark medication in the same dose regimen corresponding to 240 mg salicin/day (manuscript in preparation). As expected, we were able to detect the appearance of salicylic acid in the serum, with a peak level of approximately 1.4 mg/L. In contrast, peak levels of 35–50 mg/L have been reported after intake of 500 mg acetylsalicylic acid. The observed analgesic effect of willow bark may therefore not be attributed to salicylic acid alone. It can be speculated that other constituents of willow bark (e.g. tannins, flavonoids, salicin, salicin esters or others) may contribute to the overall effect.

Within the 2 weeks treatment period, the achieved effect was not very high (14% reduction of the WOMAC pain score). In comparison, diclofenac, a standard NSAID in the treatment of OA, has been reported to cause a 19% reduction of the WOMAC pain score from baseline in patients with OA of the knee, using 3×25 mg diclofenac/day for 6 weeks (Bellamy *et al.*, 1992), or a 23% reduction using 3×50 mg/day for 2 weeks (Bellamy *et al.*, 1993). Tenoxicam (20 mg/d) caused 15% reduction after 2 weeks (Bellamy *et al.*, 1993). Figure 1, as well as the results of the cited studies, suggests that the maximum treatment effect was not yet reached upon termination of our study (i.e. after 2 weeks). The final overall assessment in our study (Fig. 2) indicates an efficacy of this herbal remedy that clearly justifies a further careful investigation. An additional study, with a longer treatment duration and with the inclusion of a control group receiving a standard NSAID (diclofenac) is now in preparation in order to allow an evaluation of the clinical relevance of the analgesic effect of willow bark.

Note added in translation

The analgesic effect of willow bark extract, in the same dose as in the present study, has recently been confirmed in an independent clinical trial in Israel (Chrubasik *et al.*, 2000).

Acknowledgements

The study group at the Federseeklinik included the following medical doctors (numbers of patients investigated): E. Heckenberger (overall coordination; 8), E. Müller (31), S. Fahrner (13), A. Funk (13), E. Zoll (11), C. Horter (7), I. Auner (3), G. Lemmermeyer and T. Vilbrandt. We thank the nursing and laboratory staff of the Federseeklinik for support and cooperation, Dr Günaydin, University Hospital Tübingen, for assistance in the initiation of the study and C. Biegert for help in the preparation of the manuscript. We are grateful to B. Meier, Zeller AG, Switzerland, to Salushaus GmbH, Germany, and to Robugen GmbH, Germany, for their technical help in this project. The study was financed by a grant from the Alfried Krupp von Bohlen und Halbach Foundation (to L.H.), and the contribution by R.L. was funded by the Karl und Veronica Carstens Foundation.

REFERENCES

Altman RD. 1991. Criteria for classification of clinical osteoarthritis. *J Rheumatol* **18** (Suppl. 27): 10–12.

Aronson SM. 1994. A hesitant look at the future. *R I Med* **77**: 159–161.

Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Duku E. 1990. Signal measurement strategies: are they feasible and do they offer any advantage in outcome measurement in osteoarthritis. *Arthritis Rheum* **33**: 739–745.

Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. 1988. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* **15**: 1833–1840.

Bellamy N, Buchanan WW, Chalmers A *et al.* 1993. A multicenter study of tenoxicam and diclofenac in patients with osteoarthritis of the knee. *J. Rheumatol* **20**: 999–1004.

Bellamy N, Kean WF, Buchanan WW, Gerecz-Simon E, Campbell J. 1992. Double-blind randomized controlled trial of sodium mecofenamate (Meclomen) and diclofenac sodium (Voltaren): post validation reapplication of the WOMAC osteoarthritis index. *J. Rheumatol* **19**: 153–159.

Bellamy N, Kirwan J, Boers M *et al.* 1997. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* **24**: 799–802.

Blumenthal M (ed). 1998. *Complete German Commission E Monographs. Therapeutic Guide to Herbal Medicines, Developed by a Special Committee of the German Federal Institute for Drugs and Medical Devices*. American Botanical Council: Austin; Integrative Medicine Communications: Boston.

Chrubasik S, Eisenberg E, Balan E, Weinberger T, Luzzati R, Conradt C. 2000. Treatment of low back pain exacerbations with willow bark extract: a randomized double-blind study. *Am J Med* **109**: 9–14.

ESCOP. 1997. Salicis cortex – Willow Bark. In *ESCOP Monographs*; Fascicule 4. European Scientific Cooperative on Phytotherapy: Exeter.

Hochberg MC, Altman RD, Brandt KD, Moskowitz RW. 1997. Design and conduct of clinical trials in osteoarthritis: preliminary recommendations from a task force of the Osteoarthritis Research Society. *J Rheumatol* **24**: 792–794.

Julkunen-Tiitto R. 1986. A chemotaxonomic survey of phenolics in leaves of northern Salicaceae species. *Phytochemistry* **25**: 663–667.

Lequesne M. 1991. Indices of severity and disease activity for osteoarthritis. *Semin Arthritis Rheum* **20**: 48–54.

Mankin HJ. 1989. Clinical features of osteoarthritis. In *Textbook of Rheumatology*, 3rd edn. Kelley WN, Harris ED, Ruddy S, Sledge CB (eds). WB Saunders: Philadelphia; 1420

Mayer RA, Mayer M. 1949. Biologische Salicylattherapie mit Cortex salicis. *Pharmazie* **4**: 77–81.

Meier B, Lehmann D, Sticher O, Bettschart A. 1985. Identifikation und Bestimmung von je acht Phenolglykosiden in Salix purpurea und Salix daphnoides mit moderner HPLC. *Pharm Acta Helv* **60**: 269–275.

Meier B, Sticher O, Bettschart A. 1985. Weidenrinden-Qualität: Gesamtsalicinbestimmung in Weidenrinden und Weidenrindenpräparaten mit HPLC. *Deutsche Apotheker-Zeitung* **125**: 341–347.

Mills SY, Jacoby RK, Chacksfield M, Willoughby M. 1996. Effect of a proprietary herbal medicine on the relief of chronic arthritic pain: a double-blind study. *Br J Rheumatol* **35**: 874–878.

Pentz R, Busse HG, König R, Siegers CP. 1989. Bioverfügbarkeit von Salicylsäure und Coffein aus einem phytoanalgetischen Kombinationspräparat. *Z. Phytother* **10**: 92–96.

Schaffner W. 1997. Weidenrinde – Ein Antirheumatikum der modernen Phytotherapie? In *Rheumatherapie mit Phytopharmaka*, Chrubasik S, Wink M (eds). Hippokrates: Stuttgart, Germany; 125–127.

Stone E. 1763. An account of the success of the bark of the willow in the cure of agues. *Phil T Roy Soc B* **53**: 195–200.

Steinegger E, Hövel H. 1972. Analytische und biologische Untersuchungen an Salicaceen-Wirkstoffen, insbesondere an Salicin: I. Identifizierungs-, Isolierungs- und Bestimmungsmethoden. *Pharm Acta Helv* **47**: 133–141.

Stucki G, Meier D, Michel BA, Tyndall AG, Dick W, Theiler R. 1996. Evaluation einer deutschen Version des WOMAC. *Z Rheumatol* **55**: 40–49.

Weyers W. 1992. Phytotherapie in der Behandlung extraartikulärer Rheumaerkrankungen. *Therapiewoche Schweiz* **8**: 832–838.