Rheumatology 2001;40:1388–1393

Treatment of low back pain with a herbal or synthetic anti-rheumatic: a randomized controlled study. Willow bark extract for low back pain

S. Chrubasik[1,2], O. Künzel[2], A. Model[2], C. Conradt[3] and A. Black[4]

[1]*Herbal Medicines Research and Education Centre, University of Sydney, Sydney, NSW 2006, Australia,* [2]*Department of Forensic Medicine, University of Freiburg, Albertstrasse 9, 79104 Freiburg,* [3]*Department of Medical Biometry, Im Neuenheimer Feld 305, University of Heidelberg, 69120 Heidelberg, Germany and* [4]*Department of Anaesthesia, University of Bristol, Bristol Royal Infirmary, Bristol BS2 8HW, UK*

Abstract

Objectives. To compare the effects of a proprietary extract of willow bark (Assalix) and a selective inhibitor (rofecoxib) of the enzyme cyclo-oxygenase-2 (COX-2).

Methods. An open, randomized, post-marketing study was carried out in an out-patients clinic on two groups of patients aged 18 to 80 yr presenting over a 6-month period with acute exacerbations of low back pain. Using computer-generated random list, 114 patients were allocated to receive a daily dose of herbal extract containing 240 mg of salicin [PAID (phyto-anti-inflammatory drug) group] and 114 were allocated to receive 12.5 mg of the synthetic COX-2 inhibitor rofecoxib [NSAID (non-steroidal anti-inflammatory drug) group]. The doses were chosen according to existing recommendations. All patients were free to use whatever additional conventional treatments were thought necessary. The outcome measures were a modified Arhus index, its pain component and the Total Pain Index.

Results. Groups were well matched. After 4 weeks of treatment, the Arhus index had improved by about 20%, its pain component by about 30% and the Total Pain Index by about 35%. The number of pain-free patients (visual analogue scale score <2) was about 20 in each group. About 60% of the patients in each group responded well to the treatment (as judged by an improvement of $\geqslant 30\%$ in the Total Pain Index relative to its baseline). The improvement was also reflected reasonably well in the physicians' and patients' judgements of the effectiveness of treatment, which were largely concordant. Few patients of either group resorted to the additional conventional treatment options. The incidence of adverse events was similar in the two groups. Treatment with rofecoxib was about 40% more expensive than that with Assalix.

Conclusion. There was no significant difference in effectiveness between the two treatments at the doses chosen. Treatment with Assalix was less expensive.

KEY WORDS: Low back pain, Willow bark, COX-2 inhibitor, Randomized controlled study.

In several countries, the published clinical guidelines for the management of low back recommend regular treatment with analgesics and/or non-steroidal anti-inflammatory drugs (NSAIDs) [1]. These drugs are undoubtedly more effective than placebo [2] but the risk of adverse side-effects [3, 4] prompted the introduction of a new class of NSAIDS, the selective COX-2 inhibitors. Unlike the non-selective NSAIDs, these drugs produce potent analgesia with a significantly lower risk of gastrointestinal toxicity [5]. A recent study that was compliant with good clinical practice showed that a proprietary willow bark extract (Assalix) containing 15% salicin (the marker that is used for the purposes of standardization of willow-bark extracts) alleviates low back pain with no specific adverse events except occasional allergy [6]. The present study was carried out in order to compare directly the effectiveness of this phyto-anti-inflammatory drug (PAID) in treating low back pain with a representative NSAID that selectively inhibits COX-2. We chose rofecoxib as the representative synthetic

Submitted 30 December 2000; revised version accepted 4 June 2001.

Correspondence to: S. Chrubasik, Department of Forensic Medicine, Albertstrasse 9, 79104 Freiburg, Germany.

COX-2 inhibitor because, at the time the study started, there was no other COX-2 inhibitor on the German market.

Methods

Design of the study

We carried out the comparison as an open randomized study that contributed to the post-marketing surveillance of both Assalix and rofecoxib. It was conducted between January and June 2000 in an out-patient clinic in Freiburg where both treatments are prescribed routinely in the everyday management of low back pain. The principal study interventions were to use them on suitable consecutive patients according to a predetermined computer-generated random sequence and to record the effects formally using standard instruments. This was explained to each patient, along with the allocated treatment option, as part of the process of obtaining informed consent. The study proposals were approved by the ethics committee of the University of Freiburg.

Selection of patients

The patients' eligibility to participate was checked by one of the authors and an additional senior physician. The patients had to be aged between 18 and 80 yr and to have had at least 6 months of susceptibility to low back pain that was not attributable to any identifiable cause, such as disc prolapse, spondylolisthesis, osteomalacia or inflammatory arthritis. Specific exclusion criteria included: any recent trauma (because of the possibility of fracture); age >50 or <20 yr, a history of cancer or risk factors for spinal infection (recent bacterial infection; intravenous drug abuse or immune suppression); constitutional symptoms such as unexplained weight loss or recent fever or chills; or pain exacerbated by being supine or severe nocturnal pain; perineal anaesthesia; recent onset of bladder dysfunction or severe or progressive neurological deficit in the lower extremity (as a possible indication of cauda equina syndrome). Conventional generic exclusion criteria were also applied: current or recent participation in any other clinical study; serious organic illness affecting any organ system; a history of drug or alcohol abuse or requirement for psychotherapeutic agents; pregnancy or lactation; known allergy to salicylates; difficulties with language or expected cooperation.

Eligible patients were given written information sheets summarizing the treatment guidelines for low back pain, the place of willow bark extract as an established treatment in Europe, the incidence of possible side-effects of Assalix and rofecoxib, the objective and plan of the study (including the randomized allocation of treatments) and the outcome measures. The voluntary nature of participation was emphasized, as was the fact that the study data would be anonymized and protected.

Recruitment

A total of 228 patients were enrolled with informed written consent for a 4-week course of treatment. Group allocation was concealed at the time of enrolment. After completing a battery of baseline assessments (see below) and in accordance with the predetermined, computer-generated random sequence (see above), 114 patients (PAID group) were prescribed four capsules per day of Assalix (providing, *inter alia*, a daily dose of 240 mg of salicin at a daily cost of 1.24 euros) and 114 patients (NSAID group) were prescribed a single 12.5 mg tablet of the COX-2 inhibitor rofecoxib per day (at a daily cost of 1.73 euros). Participation in the surveillance did not prevent subjects from continuing with whatever other medication they usually used in the event of severe pain, or from resorting, if necessary to other conventional treatments, such as NSAIDs, acupuncture, transcutaneous electrical nerve stimulation, massage and/or other physical therapy.

Assessments

Before treatment started and 4 weeks later, the patients were assessed by questioning and examination in order to complete the standard instruments we use routinely for the documentation of the outcome of treatment for low back pain. These comprise a simple record of pain on the visual analogue scale (VAS), the modified Arhus index (mAI), its pain component (mAI-P) and the Total Pain Index (TPI) [6]. Between the start and end of the course of treatment, the patients were telephoned weekly to document any additional treatments and the occurrence of any adverse events. At the end of treatment, the physicians and patients rated the success and the acceptability of treatment on a verbal scale (very good; good; moderate; poor). Reports of adverse events were scrutinized by an independent investigator who, blinded to the allocation, rated the likelihood that the events were associated with the PAID or NSAID treatment.

Statistical analysis

This was carried out with the procedures available in Statistical Analysis System software package (SAS Institute, Cary, NC, USA). Categorical data were examined in contingency tables, with inferential testing by Fisher's exact test. Ordinal or interval data were summarized as median and quartiles (Q25;Q75) and the Mann–Whitney–Wilcoxon test was used to test differences between groups.

The percentage changes (%Δ, relative to baseline values) in the VAS, mAI, mAI-P and TPI were examined as measures of effect (using the Mann–Whitney–Wilcoxon test because the distribution has an upper bound). The principal outcome measure for the purpose of calculating study size was the percentage change in the mAI [$\%\Delta\text{mAI} = 100 \times (\text{mAI}_{\text{beginning}} - \text{mAI}_{\text{end}})/(\text{mAI})_{\text{beginning}}$]. The percentage changes in the other outcome measures were calculated in the same way. The total of 228 patients allowed the detection of a between-group difference of 0.5 S.D. in this measure

when a two-sided Wilcoxon rank sum test was used with $\alpha = 0.05$ and a power $(1 - \beta)$ of 95%.

To detect possible confounding influences on the treatment effects, linear multiple regressions were undertaken on the absolute changes in VAS, mAI-P and TPI, using a dummy variable to distinguish treatment groups, along with the fixed covariates, age, baseline value, duration of the acute exacerbation and radiation of pain into one or both legs. The confidence level for rejecting null hypotheses was taken as 95% ($P < 0.05$). The agreement between physicians' and patients' rankings of treatment in individual cases was tested with Spearman's rank correlation. Compliance with treatment was examined formally by comprehensive enquiry about each patient's consumption of medications.

Results

Table 1 lists, by treatment group, the baseline characteristics of the 228 patients who entered the study. The groups were reasonably similar in age, sex, height, weight and the duration and severity of pain, though patients in the NSAID group tended to be slightly younger, in slightly more pain and more likely to have pain radiating into one or both legs. One hundred and eighty-three patients completed the study. Table 1 lists the reasons for withdrawal in the 45 patients who did so. Non-compliance included failure to start treatment (PAID, 2; NSAID, 1), failure to attend final examination (PAID, 2; NSAID, 2), holidays (PAID, 1; NSAID, 1) and trauma (PAID, 0; NSAID, 2). Table 2 details the adverse events that occurred in 50 patients and the physicians' judgements of the likelihood that they were related to the PAID or NSAID. Further details are available at http//www.ukl.uni-freiburg.de/rechtmed/salix-rofecoxib.html.

Analgesic effects

Table 3 summarizes the outcome measures %ΔVAS, %ΔmAI-P, %ΔmAI and %ΔTPI over the 4 weeks of treatment. Irrespective of treatment group, the VAS had improved by about 44%, mAI-P by about 30%, mAI by 21–22% and the TPI by 34–35%. The number of patients with a VAS score below 2 (considered to be pain-free) at the end of 4 weeks was 22 in the PAID group and 20 in the NSAID group. Table 3 also summarizes the extent of reliance on additional treatments. In the 21 patients who resorted to additional NSAIDS, the average requirement for NSAIDS over 4 weeks in the PAID group was 120 mg diclofenac equivalents (800 mg iboprufen was deemed equivalent 100 mg diclofenac) and 5 mg tramadol, and in the NSAID group it was 42 mg diclofenac equivalent and 17 mg tramadol.

Table 4 summarizes the multivariable analyses of the changes in mAI-P and TPI, neither of which identified any significant difference related to PAID *vs* NSAID or any other significant covariable, except that a larger baseline value of mAI-P was associated with a larger change in that value. Tables 5 and 6 summarize the physicians' and patients' judgements of the effectiveness and acceptability of treatment. The ratings of the physician and patient correlated well in individual cases

Table 1. Baseline characteristics of the patients in the two groups receiving PAID and NSAID

	PAID group			NSAID group		
	n (%)	Median	Q25;Q75	*n* (%)	Median	Q25;Q75
Total number of patients in group	114 (100)			114 (100)		
No. of males	42 (37)			40 (35)		
Age (yr)		63	55;71		59	50;66
Height (cm)		168	162;173		168	164;175
Weight (kg)		77	66;84		76	67;85
Duration of low back pain						
Numbers with						
susceptibility >6 yr	87 (76)			85 (75)		
acute, <1 week	7 (6)			5 (4)		
acute, >1 week	40 (35)			53 (46)		
acute, >3 months	67 (59)			56 (49)		
Number with radiation of pain into leg(s)	35 (31)			52 (46)		
Severity of low back pain						
VAS		5	4, 7		6	5, 7
Components of modified AI						
mAI-P		22	18, 35		26	19, 39
Invalidity index		15	12, 20		17	12, 20
Physical impairment		18	16, 22		18	16, 20
Total modified AI		59	49, 70		63	51, 76
TPI		26	20, 31		29	22, 34
Number who did not complete the study because of	21 (18)			24 (21)		
non-compliance	5			6		
severe low backpain	1			3		
other pain syndromes or conditions				3		
adverse events	12			14		

Table 2. Adverse events with physicians' judgement of likelihood of causal connection with drug in treatment group

	Group	Clearly none	Unlikely	Possible	Likely	Clear connection	Withdrawals from study
Gastrointestinal complaints	PAID		2[a]	7[b]	3[c]	1[d]	4
	NSAID	1[e]		8[f]	7[g]	1[h]	9
Allergy	PAID			1	3	1	4
	NSAID						
Asthma	PAID						
	NSAID					1	1
Dizziness	PAID		1	1			1
	NSAID	1		3	1		2
Headache	PAID			1			
	NSAID			2			1
Oedema	PAID						
	NSAID			1			1
Blood pressure instability	PAID			1			1
	NSAID				1		
Sensation of heat	PAID		1				1
	NSAID						

[a]Very mild dyspepsia.
[b]Mild dyspepsia 4, 1 hiccup, 1 heartburn, 1 constipation.
[c]Dyspepsia, vomiting, heartburn.
[d]Diarrhoea.
[e]Emesis.
[f]Dyspepsia 4, abdominal pain 1, heartburn 1, flatulence 1, nausea 1.
[g]Abdominal cramps 2, abdominal pain 1, dyspepsia 2, peptic ulcer 1, nausea 1.
[h]Gastrointestinal bleeding.

Table 3. Change in pain indices and use of additional treatments

	PAID group (93 patients)	NSAID group (90 patients)	*P*
Percentage change in index [median (Q25;Q75)]			
VAS	44 (8;74)	44 (9;67)	0.94
mAI-P	30 (9;51)	32 (7;56)	0.78
mAI	21 (9;33)	22 (6;41)	0.76
TPI	35 (10;62)	34 (7;72)	0.41
Use of additional treatments [*n* (%)]			
NSAIDS and/or tramadol	9 (10)	12 (13)	0.44
Other treatments, alone or in combination	13 (14)	17 (19)	0.56
Exercises	10	11	
Physical therapy[a]	4	8	
Other[b]	3	2	

[a]Massage, heat, spa, stretching, electrical therapy, manipulation.
[b]Centrally acting muscle relaxants and analgesics, steroids, other herbal medicines, local anaesthetics, magnetic field therapy, reflex zone therapy, acupuncture.

Table 4. Results of multivariable modelling to examine for possible covariates of treatment effect

Explanator	Regression coefficient	Standard error of regression coefficient	*P*
Ordinary multiple regression on change in mAI-P			
Radiation into legs (yes/no)	−12.1	9.5	0.20
Duration of acute exacerbation (>3 months)	−8.1	5.2	0.12
Baseline AI-P	1.17	0.40	<0.01
Age (yr)	−0.36	0.20	0.08
NSAID *vs* PAID	−5.2	5.2	0.31
Ordinary multiple regression on change of TPI			
Radiation into legs (yes/no)	6.6	6.3	0.30
Duration of acute exacerbation (>3 months)	−9.7	6.2	0.12
Baseline TPI	0.49	0.39	0.22
Age (yr)	−0.27	0.24	0.26
NSAID *vs* PAID	3.0	6.2	0.62

Table 5. Physicians' and patients' ratings of effectiveness

	PAID group	NSAID group	Total
Physicians' assessment			
Very good	11 (11.8%)	15 (16.7%)	26 (14.2%)
Good	43 (46.2%)	42 (46.7%)	85 (46.5%)
Moderate	37 (39.8%)	23 (25.6%)	60 (32.8%)
Poor	2 (2.2%)	10 (11.1%)	12 (6.6%)
Total	93 (100%)	90 (100%)	183 (100%)
Patients' assessment			
Very good	9 (9.7%)	15 (16.7%)	24 (13.1%)
Good	44 (47.3%)	44 (48.9%)	88 (48.1%)
Moderate	37 (39.8%)	19 (21.1%)	56 (30.6%)
Poor	3 (3.2%)	12 (13.3%)	15 (8.2%)
Total	93 (100%)	90 (100%)	183 (100%)

Table 6. Physicians' and patients' ratings of acceptability

	PAID group	NSAID group	Total
Physicians' assessment			
Very good	33 (35.5%)	28 (31.1%)	61 (33.3%)
Good	54 (58.1%)	55 (61.1%)	109 (59.6%)
Moderate	5 (5.4%)	6 (6.7%)	11 (6.0%)
Poor	1 (1.1%)	1 (1.1%)	2 (1.1%)
Total	93 (100%)	90 (100%)	183 (100%)
Patients' assessment			
Very good	31 (33.3%)	29 (32.2%)	60 (32.8%)
Good	55 (59.1%)	57 (63.3%)	112 (61.2%)
Moderate	6 (6.5%)	3 (3.3%)	9 (4.9%)
Poor	1 (1.1%)	1 (1.1%)	2 (1.1%)
Total	93 (100%)	90 (100%)	183 (100%)

(Spearman's $r = 0.929$ for effectiveness and $r = 0.976$ for acceptability).

Adverse effects

There were 23 adverse events in the PAID group and 27 in the NSAID group. Table 2 gives the independent scrutineer's ratings of the likelihood of association for the two principal treatments. There were 30 gastrointestinal side-effects (PAID group, 13; NSAID group, 17); they tended to be more severe in the NSAID group (ulcer, 1; gastrointestinal bleeding, 1) and caused more withdrawals from the study (NSAID group, 9; PAID group, 4). There were five cases of cutaneous allergy in the PAID group and one case of asthma in the NSAID group.

Discussion

A recent systematic review of 19 double-blind, placebo-controlled randomized trials suggested that several PAIDs have some potential for alleviating rheumatic pain [7]. Two studies compliant with good clinical practice showed that Assalix was more effective than placebo in alleviating osteoarthritic pain [6, 8]. One showed a difference in effectiveness between daily doses containing 120 and 240 mg of the marker salicin [6]. A dose-dependent effect was also seen in a non-randomized open comparison of the safety and economics of treatment with Assalix with those of the treatment offered by specialist orthopaedists within the allowances and constraints of their budgets [9]. The authors calculated that the cost of treating low back pain could be reduced by including willow bark extract in the overall treatment strategy, by reducing reliance on the more expensive conventional treatments. It was argued that the difference might not necessarily have remained if the orthopaedists had placed heavier reliance on conventional analgesics and NSAIDs in their treatment strategy, though it was conceded that a different price may have been exacted in terms of the known side-effects of the conventional NSAIDS, which are non-selective in their inhibition of the enzyme COX-2. The selective COX-2 inhibitors that are now more widely available should avoid the worst of these effects, but they are also far more expensive than the non-selective NSAIDs. It was reasonable, therefore, to speculate whether there would be any clinically important difference in effectiveness between a representative PAID (Assalix) and a representative COX-2 selective NSAID (rofecoxib) if both were used at the maximally effective dose. No information is yet available on any dose-equivalence between Assalix and rofecoxib. The doses compared in this study were simply the recommended doses in the available publications [6, 9–11].

However, despite this and the imperfections in the design of this study, our present results provide some preliminary indication of the range of likely dose-equivalence between Assalix and rofecoxib. No significant difference ($\alpha = 0.05$) could be detected in a study with 95% power to detect a difference as small as 0.5 S.D. of the principal outcome measure (i.e. 10% in %ΔmAI). Though reasonable precautions were taken to detect and allow for possible failure to detect non-equivalence because of confounding, the openness of the study means that non-equivalence might still have been hidden by bias arising from the knowledge and prejudices of the patients and physicians about the relative merits of the two treatments being compared. There may be some small difference between the treatments that would be demonstrable by the use of a larger sample size. Any such small difference that there might be in favour of rofecoxib ought to be viewed in the light of the difference in the cost of treatment.

In the global assessment of effectiveness by patients, 40% rated the treatment as only moderate or poor. This patient-centred outcome measure was encouragingly consistent with the corresponding measure provided by the physicians and also with the percentage improvement in the TPI and may be worth using in further studies to record and, if possible, identify explanators of non-response, such as expectations, education, socio-economic group etc, that may not have been adequately considered so far. The number of patients using other treatment options was small and similar across groups and certainly did not affect the impact of the PAID and NSAID treatment.

From previous open non-randomized studies, the incidence of adverse events was calculated as being about 4% [12]. In recent studies [6, 8, 9] including the present study, 15 of 520 patients (about 3%) suffered from allergic skin reactions (which disappeared soon after stopping treatment) and had an incidence of other adverse events of 11%. In the placebo-controlled studies [6, 8], there were 20 adverse events in 179 willow bark patients (11%) compared with 35 adverse events in 109 placebo patients. Field studies with several thousands of patients are required to answer the issue of safety, but reasonably rigorous post-marketing surveillance studies, such as the present one, are a reasonable start.

If the natural COX-2 inhibitor and synthetic COX-2-selective NSAIDs can indeed be confidently demonstrated to be similar in effectiveness and safety, Assalix has the current advantage of being cheaper than rofecoxib. However, the situation deserves fuller and more definitive study in the light of the recent demonstration [13] that the proprietary willow bark extract Assalix, unlike the synthetic NSAIDs, is a selective inhibitor of COX-2-mediated prostaglandin E2 release and inhibits the release of cytokines to greater or lesser degree—possibly enough to have a preventive effect on cartilage destruction [14]—which might give the willow bark extract a therapeutic advantage in addition to its lower cost.

References

1. Burton AK, Waddell G. Clinical guidelines in the management of low back pain. Bailliere's Clin Rheumatol 1998;12:17–35.
2. Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. Spine 1996;22:2128–56.
3. Raskin JB. Gastrointestinal effects of nonsteroidal anti-inflammatory therapy. Am J Med 1999;106(5B): 3–12S.
4. Whelton A. Nephrotoxicity of nonsteroidal anti-inflammatory drugs: Physiologic foundations and clinical implications. Am J Med 1999;106(5B):13–24S.
5. Langman MJ, Jensen DM, Watson DJ *et al*. Adverse upper gastrointestinal effects of rofecoxib compared with NSAIDs. J Am Med Assoc 1999;282:1929–33.
6. Chrubasik S, Eisenberg E, Balan E, Weinberger T, Conradt C. Treatment of low back pain exacerbations with willow bark extract: a randomized double-blind study. Am J Med 2000;109:9–14.
7. Ernst E, Chrubasik S. Phyto-anti-inflammatories. A systematic review of randomized, placebo-controlled, double-blind trials. Rheum Dis Clin N Am 2000;26:13–27.
8. Schmid BM. Behandlung von Cox- und Gonarthrosen mit einem Trockenextrakt aus *Salix purpurea* and *daphnoides*. PhD thesis, University of Tübingen, 1998.
9. Chrubasik S, Künzel O, Black A, Conradt C, Kerschbaumer F. Potential economic impact of using a proprietary willow bark extract in outpatient treatment of low back pain: an open non-randomized study. Phytomedicine 2001;8:241–51.
10. Day R, Morrison B, Luza A *et al*. A randomized trial of the efficacy and tolerability of the COX-2 inhibitor rofecoxib vs ibuprofen in patients with osteoarthritis. Arch Intern Med 2000;160:1781–7.
11. Cannon GW, Caldwell JR, Holt P *et al*. Rofecoxib, a specific inhibitor of cyclooxygenase 2, with clinical efficacy comparable with that of diclofenac sodium: results of a one-year, randomized, clinical trial in patients with osteoarthritis of the knee and hip. Phase III Protocol 035 Study Group. Arthritis Rheum 2000;43:978–87.
12. ESCOP Monographs. Salicis cortex. Exeter: ESCOP Monograph Binder. IISBN 1-901964-03-5.
13. Chrubasik S, Fiebich B, Black A, Pollak S. Treating low back pain with a *Salix* extract that inhibits COX-2 and the release of cytokines. Eur J Anaesthesiol Suppl; in press.
14. Goldring MB. The role of cytokines as inflammatory mediators in osteoarthritis: lessons from animal models. Connect Tissue Res 1999;40:1–11.