Study Design

The study was a double-blind, 7-d randomized controlled trial comparing placebo, rofecoxib 50 mg, and diclofenac 150 mg using a double-dummy technique. The design of the study was approved by the ethics committee of Cochin Hospital (Paris, France). All patients gave their written informed consent before entering the trial. The study was conducted in 47 centers from 15 April 2003 to 16 March 16 2004.

Interventions

Study drugs. After confirmation of patient eligibility and after written informed consent was obtained, patients were randomly assigned to receive placebo, diclofenac 50 mg three times daily, or rofecoxib 50 mg once daily. All the patients took four capsules per day (two at breakfast, one at lunch, and one at dinner) during the seven days of the trial, regardless of the level of symptoms and the randomization group. Capsules and packages were identical in appearance.

Compliance was evaluated by pill count at the final visit.

Rescue therapies. Acetaminophen (500 mg tablets, maximum eight tablets/day) was used as analgesic treatment during the study when needed. Since acetaminophen was supplied as part of the study, a pill count of acetaminophen was performed at final visit.

In cases of persistent intolerable pain, a local injection of steroids was performed. This injection was defined as treatment failure, resulting in the withdrawal of the patient from the trial.

Objectives

The objective of this trial was to demonstrate the superiority of NSAIDs over placebo in acute shoulder pain over a 7-d treatment period.

Outcomes

Primary outcome measure. Changes in pain were considered as the primary outcome measure. The original trial protocol was referring to a "clinically relevant" definition of the primary outcome (e.g., a success defined by a sustained improvement of at least 50% and an absolute level of pain of 30 or less [0–100 normalized scale]). While the recruitment of patients was still ongoing and based on discussion concerning the potential loss of statistical power by using a dichotomous variable instead of a continuous variable, an amendment was proposed and accepted by the ethical committee to redefine the primary variable as the mean changes in pain during the study. Since both techniques resulted in similar findings, we are presenting here the results according to the original trial protocol. For this purpose, a diary was provided to the patient in order to collect twice a day (in the morning and in the evening) his/her pain intensity over the 12 previous hours ("nocturnal" pain collected in the morning and "diurnal" pain collected in the evening) using a 0–10 NRS.

Secondary outcome measures. Functional impairment and patient's global assessment were considered as secondary symptomatic outcome measures. Clinical assessment was performed at baseline and after 7 d by the same investigator, and functional impairment and patient's global assessment were collected. Functional impairment was evaluated using the function subscale of Neer's index [21]. This scale consists of ten questions related to daily activities (1, use back pocket; 2, perineal care; 3, wash opposite axilla; 4, eat with utensil; 5, comb hair; 6, use hand with arm at shoulder level; 7, carry 10–15 pounds with arm at side; 8, dress; 9, sleep on side; 10, do usual work). For each question, a score of 0 was assigned if the activity could be performed without any difficulty, 1 with some difficulty, 2 with marked difficulty, 3 with great difficulty (requiring assistance), and 4 if impossible. Therefore, this scale ranges from 0 to 40.

Patients' global assessments were evaluated using three different techniques. In one, at the baseline and the final visits, patient's global assessment of disease activity was collected by the following question "considering all the ways your shoulder disease affects you, mark an (X) in the appropriate box for how well you are doing" and the following potential answers 0, very well; 1, well; 2, fair; 3, poor; and 4, very poor. For the analysis, this variable was considered as a continuous one and normalized from 0 = best condition to 100 = worst condition.

In a second technique, at the final visit, patient's global assessment on her/his relative condition was collected by the following question: "compared to when you started the study, how have you been during the last 48 hours?" and the following potential 15 answers from −7, very great deal worse to +7, very great deal better. For the analysis, we considered this variable as a dichotomous one (e.g., improvement yes/no in which an improvement was considered for the patients answering at least "good deal better").

In a third technique, at the final visit, acceptable symptom state was assessed by the following question: "considering your current level of pain and functional impairment, if you were to remain for the next following months as you were during the last 48 hours would this be acceptable or unacceptable to you?" and the potential following answers 0, acceptable; 1, not acceptable.

Moreover, the requirement for rescue therapies (e.g., acetaminophen intake and/or local injection of steroids) was also considered as a secondary outcome measure of efficacy.

At baseline and final visit, blood pressure and body weight were systematically collected and at final visit, the investigators checked for tolerability.

Sample Size

The sample size needed in order to demonstrate a statistically significant difference between NSAID and placebo using the pain intensity (mean of nocturnal and diurnal pain) was calculated. A success was a priori defined by a sustained improvement of at least 50% and an absolute level of pain 30 or less (on a 0–100 normalized scale). The analysis was conducted using the Kaplan-Meier technique in which the event was defined by the time the patient fulfilled the above definition of improvement in pain with such an improvement sustained until the end of the study. Based on information on changes in pain in previously reported trials [22], the Kaplan-Meier estimates of the percentage of patients achieving such an improvement was expected to be around 40% in the active group. Since, to our knowledge, no information was available in the literature concerning the placebo group, we a priori and arbitrarily expected a 20% success rate in the placebo group. Thus a sample size of 82 patients per treatment group would allow demonstration of this difference with an alpha level of 0.05 and power of 0.80 two-tailed).

Randomization: Sequence Generation

A computer-generated randomization sequence assigned participants in a 1:1:1 ratio to receive rofecoxib, diclofenac,