The Efficacy and Cost Effectiveness of N of 1 Studies with Diclofenac Compared to Standard Treatment with Nonsteroidal Antiinflammatory Drugs in Osteoarthritis

JANET E. POPE, MARK PRASHKER, and JENNIFER ANDERSON

ABSTRACT. ***Objective.*** To determine whether or not N of 1 trials with diclofenac/misoprostol (Arthrotec) are superior and cost-effective compared with standard treatment in osteoarthritis (OA).

Methods. We randomized subjects with OA who were "uncertain that nonsteroidal antiinflammatory drugs (NSAID) were helpful" to 2 different groups. One group received conventional treatment whereby they were told to stop their NSAID and to wait and see what happened. If necessary, treatment with other NSAID and all other usual OA treatment strategies were used. The other group received a series of crossover trials with diclofenac 50 mg and misoprostol 200 µg twice a day or an identical placebo for 2 weeks each in a random double-blinded manner. Every 4 weeks they chose which treatment they preferred. By 3 months, if there was no clear preference, the N of 1 trials were discontinued. All trial participants were seen monthly for 3 months and at 6 months. All costs (direct and indirect) were collected for both groups. Costs of research-generated visits were not counted in the "conventional treatment" group.

Results. Fifty-one subjects were randomized (stratified by most symptomatic OA area): 25 with knee, 7 with hip, and 19 with hand OA. Twenty-four were randomized into the N of 1 group. There were no differences in the baseline and followup variables including age, income, education, past and current NSAID use, global assessment, Health Assessment Questionnaire (HAQ), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), and Medical Outcome Study Short Form-36 (SF-36) scores. In the N of 1 group, 11 patients preferred diclofenac in total (of whom 7 did only one round, 2 did 2 rounds, and 2 did 3 rounds). None of the N of 1 patients preferred placebo and 11 had no preference (2 dropped out at baseline). At 6 months, 15 of 19 in "conventional therapy" and 17 of 21 in N of 1 were taking NSAID. That is, NSAID appeared to be effective in 81% of N of 1 subjects and 79% of conventionally treated patients, even though subjects were initially uncertain that their NSAID were helpful. The total OA-related costs in Canadian dollars per patient (in 1996) for N of 1 treated patients at 6 months were: $551.66 ± $154.02 (SD) versus $395.62 ± $226.87 for controls, excluding 2 research visits for controls ($p < 0.009$). The HAQ pain and disability, WOMAC scales, and physician global assessments improved more in the N of 1 group (at greater cost), but no between-group differences in efficacy were seen, possibly due to small numbers.

Conclusion. N of 1 trials were time-consuming in these patients and are more expensive, but with slightly better outcomes. In addition, NSAID seem to be effective in a majority of subjects with OA who have been uncertain of their benefit. (J Rheumatol 2004;31:140–9)

Key Indexing Terms:
N of 1 COST EFFECTIVENESS TRIAL RANDOMIZED
OSTEOARTHRITIS NONSTEROIDAL ANTIINFLAMMATORY DRUGS

Symptomatic osteoarthritis (OA) is very common, with several treatments, but in many patients the mainstay has been the use of nonsteroidal antiinflammatory drugs (NSAID). The American College of Rheumatology guidelines for the treatment of OA have been proposed[1,2] and include the use of NSAID, presumably in those who have failed treatment with acetaminophen. More recently guidelines have been published in which NSAID may be used as first-line treatment in certain patients with OA[3,4].

OA is one of the most prevalent diseases in North America. The costs of NSAID are a major factor in the care of rheumatology patients[5]. Yet 20% to 30% of patients with OA in clinical trials have no preference between active

From the Department of Medicine, Division of Rheumatology, University of Western Ontario, London, Ontario, Canada; Boston University Schools of Medicine and Public Health, Boston, Massachusetts; and the Center for Health Quality and Outcomes and Economic Research, Bedford Veteran's Administration Hospital, Bedford, Massachusetts, USA.

Supported by Physicians Services Incorporated Foundation, Toronto, Ontario.

J.E. Pope, MD, MPH, FRCPC, Associate Professor of Medicine, University of Western Ontario, St. Joseph's Health Care London; M. Prashker, MD, MPH; J. Anderson, PhD, Boston University Schools of Public Health and Medicine and the Center for Health Quality Outcomes and Economic Research.

Address reprint requests to Dr. J.E. Pope, Rheumatology Centre, St. Joseph's Health Care London, 268 Grosvenor Street, PO Box 5777, London, Ontario N6A 4V2.

Submitted July 15, 2003; revision accepted June 18, 2003.

NSAID and placebo. Therefore, there may be a chance through N of 1 trials to decrease NSAID medication use in patients with OA if the NSAID treatment strategy is not clearly preferred. This would also decrease the potential morbidity and costs accompanied by these drugs, and could lead to treating only those in whom NSAID were truly beneficial.

N of 1 trials (single-case research, N = 1 randomized controlled trials) are multiple, randomized, crossover comparisons of an active drug against its placebo in a single subject[6-9]. Important prerequisites for their use include: a disease that is chronic and stable; a treatment that starts and, once withdrawn, stops acting quickly; identification of clinically important variables likely to be affected by the treatment; and difficulty making a decision about the treatment in a particular patient from a conventional randomized controlled trial. For example, it may be that trial results cannot be generalized to the clinician's particular patient, or heterogeneity of response within such trials puts the efficacy of treatment in doubt for a particular patient[6,7].

In patients with OA who are uncertain if their NSAID have been helpful, we hypothesized that the group of patients undergoing N of 1 trials would have a better overall outcome than the group given treatment established by conventional practice. In addition, we hypothesized N of 1 trials may be cost-effective relative to conventional practice; the savings from stopping ineffective therapy sooner may exceed N of 1 trial costs arising from extra demands made of physicians, patients, research assistants, and pharmacists[9]. Neither hypothesis has been adequately tested in the treatment of arthritis in a randomized controlled fashion[9-16]. March, *et al* compared acetaminophen to diclofenac in N of 1 trials in 25 patients with OA. The dropout rate was 40%; half withdrew early but did make a therapeutic decision, and the other half dropped out very early for other reasons. In this study the majority of patients had adequate symptom control using acetaminophen alone. Not all patients were taking NSAID and 60% had changed their drug from the baseline. The investigators also concluded that N of 1 trials should prove useful as a means of withdrawing NSAID from patients who have OA and may not need them any longer. However, within their trial, many patients were not withdrawn from NSAID[17]. Pincus, *et al* found 80% of OA subjects surveyed named an NSAID as most helpful compared to 20% who named acetaminophen[18]. Wolfe, *et al* also found patients with OA were more likely to prefer NSAID than acetaminophen in a mailed survey[19]. In a prospective US study it was found that in the care of OA, medications accounted for 32% of the cost[20]. Thus, if an N of 1 crossover trial could demedicate patients who truly no longer need an NSAID, this would reduce costs.

The null hypothesis of our trial was that there would be no difference in the outcomes and costs between patients with OA of the hip, knee, or hand who currently are not convinced that NSAID have been helpful, whether they are randomized to standard care or to a trial with diclofenac and placebo in a series of double-blinded N of 1 trials. We hypothesized that N of 1 trials may be cost-effective relative to conventional practice, with equal or better outcome measurements.

MATERIALS AND METHODS

This was a parallel trial, where subjects who were not certain that NSAID were helpful or were unexposed to NSAID were randomized to clinical practice ("conventional treatment") or to N of 1 trials, and then followed for 6 months. All subjects were stratified by former use of NSAID (yes or no) and by the most symptomatic area of OA (hip, knee, or hands). The N of 1 trials were double-blinded. Subjects received 2 weeks of diclofenac (Arthrotec®; supplied by Pharmacia-Searle Canada, Toronto, ON) or placebo in random order, and continued for a maximum of 3 cycles (3 months). A decision was made to continue or stop the study drug based upon the patient's global assessment in each of the treatment periods (final day of each 2-week treatment). Figure 1 illustrates how the trial was carried out, showing only one of the 3 strata. Final followup for cost comparison was at 6 months after entry, with baseline measurements being repeated at 3 and 6 months. We extended followup 3 months beyond the trial to determine if the rates of NSAID use and OA utilization differed between the 2 groups over time. To balance the potential bias of frequent outcome measurements and visits to the research personnel in the N of 1 group, both groups were seen and completed global assessments: the Stanford Health Assessment Questionnaire (HAQ)[21] and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[22] (if they did not have hand OA), monthly for 3 months and once again at 6 months. All patients were permitted analgesics as needed and the drug dosage and frequency of use were monitored. All subjects in the standard care group were recommended to stop their current NSAID if they were taking one, because they were not sure it was working, and to telephone if they worsened prior to the next followup visit. This was done to decrease the potential bias of prescribing many NSAID in this group. The research assistant was blinded to group allocation (conventional or N of 1) so the questionnaires would be administered in the same way to both groups, but the investigator and subjects were aware of group allocation. In the N of 1 group, the subject and investigator were blinded to treatment order (diclofenac and misoprostol or placebo in random order for 2 weeks each in every series).

Inclusion criteria were: (1) outpatients aged ≥ 18 years with symptomatic OA of the knee, hip, or hands with pain; < 30 minutes of morning stiffness with no evidence of other rheumatic diseases such as rheumatoid arthritis or chondrocalcinosis; (2) uncertainty in the subject's mind that current NSAID therapy was helpful (i.e., patients who answered no to the question, "Are you certain that this nonsteroidal antiinflammatory drug is helping you?" or yes to "I have not tried NSAID." If a patient had more than one area of OA, the most symptomatic area (either the hips, knees, or hands) was used to stratify the patient and this area was used for the primary outcome measurements. Exclusion criteria were as follows: (1) those with a relative contraindication to NSAID, including uncontrolled hypertension, documented peptic ulcer disease in the last 2 years, renal insufficiency with creatinine > 200 mg/dl, unexplained anemia or serious hepatic disease, and those taking anticoagulants; (2) use of oral steroids; (3) those thought likely to require joint surgery in the next 6 months; (4) women with child-bearing potential who were pregnant, breastfeeding, or not practicing contraception; (5) those with major comorbidity in whom it was believed that they would be unable to complete the trial; (6) those unable to give informed consent; (7) those with an allergy to any NSAID or aspirin; (8) those who did not tolerate diclofenac or Arthrotec if previously exposed; and (9) those who had received an intraarticular injection over the previous 3 months.

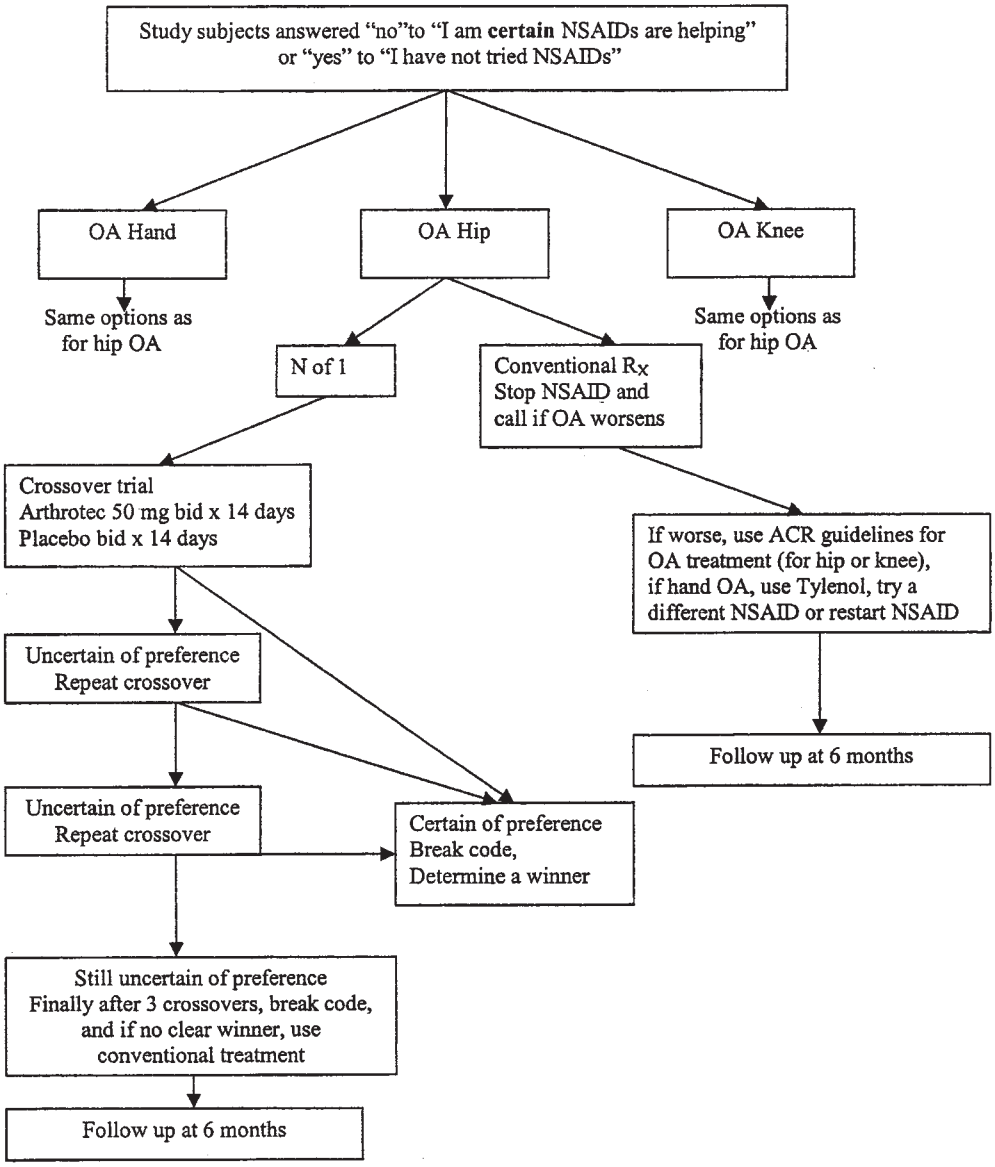

Figure 1. The trial design (hip stratum only).

There are no standardized power analyses for economic studies, so we chose to use sample sizes required for clinical measures to show differences[23]. We hypothesized our baseline population would be mild to moderate with respect to pain at baseline. To find > 20% difference between groups one would need about 26 patients per group, based on a 14 mm improvement in patient global assessment and a baseline HAQ of 0.9, a standard error of the mean (SEM) of 0.13 with 80% power, and a 2-tailed t test at $p < 0.05$. Subjects were recruited from rheumatology practices and from a newspaper advertisement. The latter strategy yielded the most patients. There were about 3 screen failures for every eligible subject, mostly because subjects said NSAID were helpful.

"Conventional treatment" group. Patients were reviewed monthly by the research assistant and physician for a period of 3 months and again at 6 months, in order to maintain concealment of treatment allocation for the research assistant, and to balance the "placebo effect" of frequent visits between the 2 treatment groups. The nurse, if unblinded, may have unwittingly biased the subject's self-assessed questionnaires (verbally or nonverbally). The outcome measurements were done at each visit. These patients received standard care by their own physician and the study investigator following guidelines for OA treatment[1-4], which could include use of high doses of acetaminophen, switching strategies to other NSAID, or other treatments such as intraarticular steroids, physiotherapy and occupational therapy. All cointerventions were recorded. At 3 and 6 months, the number of patients taking NSAID was calculated.

We artificially increased the frequency of visits for the first 3 months in the "conventional treatment" group, so the added costs of research visits in this group were ignored in the primary cost analysis.

N of 1 group. Active and placebo preparations of diclofenac were to be taken as one pill twice a day with food, in a randomized, concealed allocation in balanced blocks of 2. Subjects (and the investigator) were blinded to treatment order. Each subject was assigned a number, for which 3 sets of drugs had been predetermined. In each 2-week block of 4 weeks, Arthrotec and placebo were supplied in bottles A and B by the pharmacist. Thus, for each cycle, both the investigator and the subject were unaware of the order

in which the active drug and the placebo would be received. Each treatment pair lasted 28 days (2 weeks of drug A and 2 weeks of drug B) without washouts. Pill counts were performed at each visit.

Patient global assessments (100 mm visual analog scale, VAS) were done on the last day of each treatment with placebo or diclofenac (i.e., roughly every 2 weeks for up to 3 months in order to eliminate carryover effects), and in the control group every month for the first 3 months. The treatment code was not broken until completion of the third treatment pair unless the patient or treating physician wanted to discontinue the N of 1 trial for any other reason, such as certainty that one drug was more effective than the other, or if an adverse reaction occurred. When the code was broken, the treating physician was given a report summarizing the subject's mean global assessment scores on the outcome measurements for placebo and NSAID therapy, including the mean difference with a Student t-distribution. Preference was thus defined as a statistical improvement on VAS for global assessment while taking one treatment compared to the other. The physician and subject then decided to continue or discontinue diclofenac. Arthrotec was used because it had a safer profile with respect to cytoprotection and was slightly cheaper than other commonly used nongeneric NSAID at the time of study enrollment. Indeed at the time of the study, it was the most commonly prescribed NSAID in Canada. Pharmacia-Searle provided the active drug. The local pharmacy provided gel coating for the active drug and the placebo (consisting of a filler). The look and the taste of the placebo were identical to active drug. After 3 months, both groups received standard care. The proportion of patients taking NSAID at 3 and 6 months was compared in the 2 groups.

The primary outcome measurement was the patient global assessment. Secondary outcome measurements included physician global assessment, WOMAC[22] score for lower extremity OA and HAQ[21] score for upper extremity OA, the HAQ[21] disability and pain scores and the Lesquesne Index[24] (for patients with lower extremity OA), and the Medical Outcome Study Short Form-36 (SF-36)[25]. The patients and physicians were asked if they were certain the NSAID currently used should be discontinued or continued at all patient visits where applicable.

Other medications and interventions were recorded and determined to be either related or unrelated to the OA or the NSAID treatment. Extra visits related to OA or its treatment were recorded, such as visits and treatments for side effects such as dyspepsia, rash, etc.

The trial costs were determined using society's perspective. The subjects filled out questionnaires with relevant economic information. This information was needed to be able to calculate the opportunity cost of a subject's time. For those not working, the United States Bureau of Labor Statistics calculations for the comparable cost of household work were used. Direct costs included laboratory tests (if any), medications and treatments for OA, costs of medications for side effects, and the physician's (primary care and investigator time) and nurse's time. Indirect costs included the patient's time, time of a driver if indicated, and travel time. Physician utilization of all tests and times the research nurse and physician spent with the patient were recorded at each visit. All study variables were measured in units of care that would be identical to those used in any other health care setting. In this manner the study was not institution-specific, but it was also not specific to any particular health care delivery system (assuming that all patients had equal access to care). Costs were determined by the cost of medications, the number of medication changes, average time spent per visit (patient travel time, research assistant time, investigator time), other physician visits related to OA or its treatment, and number of cointerventions reported. Costs were calculated in Canadian dollars (for the year 1996) and were not adjusted for inflation to the current Canadian dollar. The time amounts were converted to costs using the patient's salary, hourly wage, or the appropriate conversion factor for homemakers or disabled or retired individuals. The cost data were combined with the effectiveness data. Primary cost analysis utilized only the research-driven visits for the N of 1 group (i.e., at one, 2, and 3 months).

All tests of significance were 2-sided ($p < 0.05$ as statistically significant) with an initial intent-to-treat analysis, where the last available result was carried forward on all effectiveness outcome measurements. Completers only were measured in a secondary analysis.

RESULTS

Fifty-one patients were randomized within this trial and stratified with the major symptomatic joint area being hip[7], knee[25], or hand[19]. Twenty-four were randomized to the N of 1 treatment group and 27 to the "conventional treatment" group. There were no significant differences in the baseline characteristics including mean age (54 ± 2.4 years in N of 1 group, and 59 ± 2.3 years in conventional therapy). The N of 1 group had OA for 14 years versus 12.7 years in the conventional group. Ninety-six percent of the patients were Caucasian and over 60% were married. In the N of 1 treatment group the family size was slightly smaller and more owned their own dwelling. Their average household income was roughly $3500 higher than in the conventional treatment group; however, none of these were significantly different. In the conventional therapy group more patients were retired. Table 1 displays the baseline characteristics of the groups.

The baseline HAQ score was 0.84 ± 0.12 in N of 1 versus 0.92 ± 0.12 in conventional treatment. WOMAC scores were also very similar: 239.6 in N of 1 versus 233.7 in the conventional therapy group for pain, 107.2 in N of 1 versus 116.6 in the conventional group for stiffness, and 762.7 in N of 1 vs 790.3 in the conventional group for physical function. The patient global assessment was 55.6 in N of 1 and 55.8 in the conventional treatment and the physician global assessment was 35.6 versus 33.8 on a 100 mm VAS. Higher values are equivalent to a worse rating. Current NSAID use was different at time of entry into the trial, where in the N of 1 group only 46% were currently taking NSAID, versus 70% in the conventional treatment group ($p < 0.08$). Past NSAID use was reported by most subjects (92% in each group had taken NSAID for their OA). At baseline for the N of 1 group: 9 were taking no medications, 9 NSAID, 9 acetaminophen, and 3 narcotics; and in the conventional therapy group: 6 were taking no medications, one took herbal remedies, and 5 used acetaminophen, 13 NSAID and one narcotics. Table 2 shows the interventions over 6 months as well as OA related drug changes. Medications that were altered or changed, including switching NSAID, significant dose changes, and adding analgesics, were recorded.

The dropout rate was higher in the conventional treatment group than in N of 1 (8 in conventional therapy vs 3 in N of 1). Four subjects dropped out immediately after randomization (2 in each group). Most subjects withdrew due to a lack of interest once they understood that there were no novel therapies available in the trial. In the N of 1 group 11 patients preferred NSAID over the 3 months (7 of 13 who went only one round, 2 of 3 who went 2 rounds, and 2 of 6 who went all 3 rounds). No one preferred placebo and 11 had no preference. At 6 months, 17 were taking NSAID in

Table 1. Comparison of the baseline characteristics of patients receiving either conventional or N of 1 treatment. Results are presented as means ± SEM (where applicable) with 95% confidence intervals in parentheses.

Baseline Characteristics	Treatment	
	Conventional Therapy	N of 1
N	27	24
Age, yrs	59 ± 2.3 (54, 63)	54 ± 2.4 (49, 59)
Most Symptomatic OA Joint, %		
Hip	15	13
Knee	44	54
Hand	41	33
OA Disease Duration, yrs	12.7 ± 2.2 (8.3, 17.1)	14.0 ± 2.3 (9.4, 18.5)
Ethnicity, %		
Caucasian	96	96
Other	4	4
Marital status, %		
Single	7	0
Married	67	62
Divorced/separated	15	21
Widowed	11	17
Education, %		
Elementary/middle	22	12
High school	26	42
Some college/graduate	48	46
Professional degree	4	0
Total family size, n	2.33 (0.21)	2.04 (0.29)
Own the family dwelling, %	67	79
Household income, $	35,092 ± 3594	38,646 ± 3987
Job status, %		
Full time	15	34
Part time	22	12
Homemaker	18	17
Retired	38	29
Other	7	8
Patient global assessment (100 mm VAS)	55.8 ± 4.64 (46.4, 65.2)	55.6 ± 5.10 (45.4, 65.8)
HAQ Score	0.923 ± 0.12 (0.67, 1.17)	0.838 ± 0.12 (0.58, 1.10)
WOMAC score		
Pain	233.7 ± 29.1 (174.0, 293.3)	239.6 ± 29.1 (180.0, 299.2)
Stiffness	116.6 ± 14.1 (87.7, 145.4)	107.2 ± 14.6 (77.4, 137.0)
Physical function	790.3 ± 104.4 (576.8, 1003.9)	762.7 ± 107.8 (542.2, 983.3)
Physician global assessment (100 mm VAS)	33.8 ± 4.12 (25.9, 41.69)	35.6 ± 3.97 (27.1, 44.1)
SF-36 score		
Physical	32.9 ± 10.9	35.6 ± 9.7
Mental	49.8 ± 10.7	51.0 ± 10.2
Lequesne Functional Index score	12.42 ± 1.39 (9.95, 14.88)	11.85 ± 0.94 (9.47, 14.22)
Current NSAID Use, %*	70	46
Past NSAID Use, %	93	92

* Current NSAID use was higher in conventional treatment ($p < 0.08$).

the N of 1 group and 9 were not; whereas 15 were taking NSAID and 4 were not on conventional treatment. Using last observation carried forward, 17 of 24 (71%) in the N of 1 and 18 of 27 (67%) in the conventional therapy group took NSAID. The NSAID used in the N of 1 group at 6 months were: 11 taking diclofenac (7 taking Arthrotec), 3 naproxen, and one each of tenoxicam, tiaprofenic acid and ibuprofen. In the conventional treatment group at 6 months, NSAID included: naproxen in 7, diclofenac in 5 (one Arthrotec), tenoxicam in 4, ketoprofen in one, and tiaprofenic acid in one. There were more side effects during the trial in the N of 1 treatment group (Table 3). Only one of the 4 rashes was thought to be related to the NSAID.

Table 4 shows the outcome measurements for the trial. There were no significant between-group differences. Interestingly, the HAQ measure improved slightly in the N of 1 group and worsened slightly in the control group. HAQ disability was 0.7 (N of 1) versus 0.97 (conventional) ($p < 0.1$). The same trend was found for the HAQ pain rating, which was 1.51 (conventional group) versus 1.17 (N of 1 group) ($p < 0.06$). From the SF-36 results, we observed that general health improved in both groups, but more in the N

Table 2. Concomitant treatment during the 6 months. Narcotics were counted 7 times in each group, and physiotherapy, 10 times in N of 1 and 17 in conventional treatment. Two had occupational therapy over 6 months (one in each group): 2 went to a chiropractor (both in conventional); 5 had intraarticular steroids (4 conventional and one in N of 1). No one had viscoelastic supplementation or joint surgery over the 6 months. In conventional treatment at 1 month, 12 were taking NSAID, 2 months 16, 3 months 13, and at 6 months 18.

Cointervention*	Conventional Therapy, n = 27	N of 1, n = 24
Acetaminophen		
Baseline	5	9
1 mo	8	11
2 mo	9	8
3 mo	11	8
6 mo	8	9
Narcotics		
Baseline	1	3
1 mo	1	3
2 mo	2	1
3 mo	1	0
6 mo	2	0
Physiotherapy		
1 mo	1	1
2 mo	6	3
3 mo	4	2
6 mo	5	4
OA drug change		
1 mo	14	11
2 mo	3	5
3 mo	4	5
6 mo	3	3

* Number with cointervention at each time point.

Table 3. Side effects during the study: number of patients with these complications.

	Conventional Therapy	N of 1
Heartburn/dyspepsia/diarrhea/constipation	3	12
Rash	1	4
Hypertension	0	2
Edema	0	1
Epistaxis	1	0
Oral ulcers	0	2

of 1 group. Patient global scales favored N of 1, but physician global scores were identical. Cointerventions were slightly higher in the N of 1 group and medication changes were higher in the control group, with an average of 1.85 versus 1.58 in the N of 1 group. There were more toxicities and adverse events in the N of 1 treated subjects: average per subject was 1.33 (N of 1) versus 0.26 (conventional) ($p < 0.001$). The time spent with each patient was, from the perspective of physician time, virtually identical. As expected, the N of 1 strategy required more time, as the crossover design had to be explained, pills had to be

Table 4. Primary and secondary outcomes for OA in the 2 treatment groups. Results are presented as means ± SD (when applicable), where mean represents the average outcome measurement per patient over the course of the trial.

	Treatment		
	Conventional Therapy	N of 1	p
Randomized, n	27	24	
Cointerventions per patient, total excluding baseline	2.30 ± 22.16	2.25 ± 1.96	0.937
Toxicities and adverse events			
Any toxicity, n	5	18	
Average number per patient	0.26	1.33	0.001
HAQ, n	26	24	
Disability	0.97 ± 0.66	0.7 ± 0.57	0.122
Pain	1.51 ± 0.71	1.17 ± 0.56	0.064
Physician global	35.47	36.24	0.864
Patient global	50.52	42.97	0.218
SF-36, n	26	24	
Physical component	34.05	36.06	0.428
Mental component	49.15	51.04	0.489
WOMAC, n	16	15	
Pain	43.12	38.7	0.581
Stiffness	52.12	42.46	0.325
Function	44.8	36.34	0.360
Lequesne, n	13	14	
Score	12.55	11.58	0.591

counted, and statistical tests performed by the investigator. It did not increase overall average subject time; however, nurse time was increased, as the visits for the research study were only counted in the N of 1 group. Travel time was slightly higher in the control group.

The costs in the N of 1 group accounted for all research visits, whereas in the control group the extra research visits were not added. Table 5 shows the costs during the trial using the last observation carried forward and then "completers-only" data. Drug costs were marginally higher in the conventional treatment group despite less NSAID use at 6 months in completers. More physician and research time was used in the N of 1 strategy. A sensitivity analysis was not done, as the variability in estimates was large and there were no statistically significant differences in the primary outcome measurements. Costs for time and drugs over 6 months were $533.48 for controls and $565.39 for N of 1 subjects. Control subjects used more NSAID and so their drug costs were slightly higher than the N of 1 subjects ($218.80 for conventional vs $198.74 for N of 1), offsetting the greater time-related costs for the N of 1 group ($314.68 for controls vs $366.65 for N of 1).

If the protocol-generated visits were removed in the N of 1 group after a treatment decision was made, there would have been a total of 29 fewer visits. If this is taken into consideration, and the average time costs are eliminated for those visits, then the costs of patient, MD, and nurse would have totaled $265.38 in the N of 1 group compared with

Table 5. Average time, costs, and total costs, per person over 6 months for N of 1 and conventional treatment groups. Results are presented as means ± SD, where applicable, with n indicated in parentheses. Based on $ Cdn in 1996.

	Treatment		
	Conventional Therapy	N of 1	p
Drug costs, $			
Completers (n)	218.80 ± 163.64 (19)	198.74 ± 163.64 (21)	0.680
All subjects (except the 4 who dropped out at randomization) (n)	175.08 ± 163.94 (25)	193.04 ± 140.66 (22)	0.379
Time spent (all visits)			
Patient time, hours	5.9 ± 5.4	5.9 ± 4.1	
Physician time, minutes	95.6 ± 32.5	123.1 ± 34.5	0.005
Nurse time, minutes	126.4 ± 57	148.4 ± 70	
Time costs, $	(25)	(22)	
Including all visits for conventional			
Patient time	115.36 ± 169.59	101.73 ± 56.85	0.708
MD time	166.93 ± 48.50	213.26 ± 47.64	0.002
Nurse time	36.95 ± 15.58	43.63 ± 19.09	0.194
Excluding visits 1 & 2 for conventional			
Patient time	81.36 ± 126.98	101.73 ± 56.85	0.474
MD time	108.40 ± 34.29	213.26 ± 47.64	0.001
Nurse time	30.79 ± 13.20	43.63 ± 19.09	0.009
Excluding visits 1, 2, 3 for conventional			
Patient time	67.68 ± 106.22	101.73 ± 56.85	0.172
MD time	89.73 ± 25.49	213.26 ± 47.64	0.001
Nurse time	28.77 ± 13.27	43.63 ± 19.09	0.003
Time costs (completers), $	(19)	(21)	
Including all visits for conventional			
Patient time	96.69 ± 60.57	104.71 ± 56.46	0.667
MD time	177.37 ± 36.36	217.46 ± 44.44	0.004
Nurse time	40.62 ± 14.10	44.48 ± 19.13	0.476
Excluding visits 1 & 2 for conventional			
Patient time	66.65 ± 41.45	104.71 ± 56.46	0.021
MD time	113.60 ± 26.43	217.46 ± 44.44	0.001
Nurse time	33.57 ± 11.84	44.48 ± 19.13	0.039
Excluding visits 1, 2, 3 for conventional			
Patient time	54.69 ± 36.40	104.71 ± 56.46	0.002
MD time	92.54 ± 22.46	217.46 ± 44.44	0.001
Nurse time	31.32 ± 12.48	44.48 ± 19.13	0.015
Total costs (time + drugs), $			
Including all visits for conventional	494.32 ± 266.06	551.66 ± 154.02	0.3650
Excluding visits 1 & 2 for conventional	395.62 ± 226.87	551.66 ± 154.02	0.0093
Excluding visits 1, 2, 3 for conventional	361.26 ± 204.68	551.66 ± 154.02	0.0007
Total costs (time + drugs), completers, $			
Including all visits for conventional	533.48 ± 175.33	565.39 ± 143.37	0.5308
Excluding visits 1 & 2 for conventional	432.62 ± 162.79	565.39 ± 143.37	0.0092
Excluding visits 1, 2, 3 for conventional	397.39 ± 158.12	565.39 ± 143.37	0.0011

$178.55 in the conventional group. When all visit costs were assessed, total costs in the conventional therapy group were $397.35 compared to $377.29 in N of 1, making these 2 strategies equal. However, this was a post hoc analysis using all visits despite a potentially artificial schedule of visits recommended in the conventional treatment group.

For an extra $108.74 over 6 months in the N of 1 subjects, there would be less pain and disability (0.34 less on HAQ pain and 0.27 less on HAQ disability, both on a 0 to 3 scale). The ratio would not be statistically significant, given the substantial variability in both numerator and denominator components. Also, any relative improvement in disability and/or pain among the N of 1 subjects has to be seen in the context of the larger number of subjects with toxicities and adverse events.

DISCUSSION

In our study, patients used NSAID at a higher rate in the N of 1 group over the first 3 months. They had trends toward slightly greater efficacy but at higher costs due to the time and costs of visits, a research nurse, and preparation of the study medication. There has been debate around what would

be a minimally clinically relevant difference in HAQ score in OA. Some have thought 0.2 units could be relevant. We achieved roughly this difference but, with small numbers, this did not reach statistical significance. Thus, one could conclude slightly better efficacy using an N of 1 strategy compared to conventional treatment in OA, but at higher costs. It had been found that OA subjects in trials will have increased functional improvement regardless of treatment if they flare prior to randomization[26]. In order to obtain NSAID treatment in the conventional therapy group, subjects had to experience flare. This could have biased treatment against the N of 1 practice, but this strategy was still favorable, but insignificant perhaps due to a small sample size. There was a 20% dropout rate at the final visit and the differential proportion of protocol violators in the 2 groups. More subjects in conventional treatment could have dropped out, as they were not getting special trial medication.

The rates of NSAID usage at 6 months were strikingly similar in the 2 groups. It appears that NSAID were effective in patients with OA who at baseline were not certain of their effectiveness. In this group, a simple strategy may be to stop the NSAID, and if the patient worsens, restart or switch to a different NSAID; this strategy is used routinely in clinical practice. Pincus, *et al* performed a randomized crossover trial of diclofenac versus acetaminophen in OA of the hip and knees. Diclofenac was found to be superior to acetaminophen[27]. The aim of our study was to treat NSAID nonresponders. In our subjects, many were responders, as judged by their NSAID utilization.

We may have biased the trial, as all patients were asked to discontinue their NSAID because they were not sure it was working in both the N of 1 and conventional treatment groups. The N of 1 subjects immediately received a series of diclofenac trials, whereas the conventional therapy group had no NSAID treatment, but telephoned if they worsened before the next followup visit. The initial efficacy results were biased against the conventional therapy group, as all were no longer taking NSAID, at least for a short period of time. However, we thought this was equivalent to the conservative-cost approach used in clinical practice. As this trial was carried out by a rheumatologist, many treatment approaches for OA were utilized (acetaminophen, intraarticular steroid injection, and physiotherapy). There are both advantages to one treating physician (standardized care) and disadvantages (less variability).

The positive effects of being in a trial (several visits with the physician and study nurse) were equally distributed in both groups, but the costs only applied to the N of 1 group, tending to decrease costs for increased benefit in the conventional group (that is, favorably biasing the cost-benefits in the control group). As well, in the N of 1 group, the global assessments were done at each monthly study visit, during which the subjects were potentially taking placebo half the time, until each subject clearly preferred a treatment. This would bias against efficacy over the first 3 months in the N of 1 group.

The sample size may not have been large enough to determine the differences in cost for the rare but costly complications of NSAID use. However, by 6 months the NSAID use was similar within the 2 treatment strategies, and indeed more subjects in the N of 1 group were receiving NSAID. Therefore NSAID complications could be projected to be more numerous in the N of 1 group in the long run. Indeed, in our trial there were more adverse events in the N of 1 group, but no serious adverse events such as clinically diagnosed ulcers occurred in either group. Interestingly, some added benefits seemed to occur when patients were in the N of 1 trial as measured by changes in the HAQ and some features of the SF-36, although these were not statistically significant. Our sample size may have been too small to detect differences between the 2 strategies that were statistically significant. Others have found that the WOMAC, SF-36, and HAQ provide different measures in OA of the knee[28]. The SF-36 provides general health measures and is responsive to changes in patients with OA. The Australian/Canadian Osteoarthritis Hand Index (AUSCAN), an OA-specific hand measure, was not available at the time of study design[29].

We chose 2 weeks of each treatment, as in clinical practice one should know if an NSAID is helpful before 2 weeks' duration, and subjects would have reached a steady-state by that point. We chose no washout, as this would increase the complexity of the trial (possibly resulting in further telephone calls or noncompliance), realizing that there could be carryover effects. However, within a few days an NSAID should be washed out, but we do not know if the subject returned to a steady-state.

We were not able to conceal from subjects and physicians the allocation to N of 1 trials or conventional practice, resulting in a potential bias when using subjective response variables. However, N of 1 subjects and investigators were blinded to treatment order of diclofenac and placebo. After each 4-week series of 2 weeks of taking diclofenac or placebo in random order, the code was not broken until a clear preference was found. Thus, subjects and investigators did not know the order of treatment or which drug was preferred until all rounds were done or a clear preference was made. In order to minimize bias in subject response that may arise from more intense followup during the first 3 months in the N of 1 trial group, subjects in the conventional practice arm were seen in the clinic more frequently than would normally occur under standard care.

It is likely that using a coxib in the N of 1 treatment group would have yielded similar results and costs in Canada. Nongeneric NSAID including coxibs are priced similarly in Canada. Their risk reduction in gastrointestinal (GI) events should be similar or superior to Arthrotec.

However, side effects such as diarrhea and dyspepsia could potentially be less with the coxibs, as shown in many coxib trials (slightly less dyspepsia, GI upset, and other nuisance side effects). Adding a proton pump inhibitor (PPI) to traditional NSAID (especially nongeneric) would be more costly, as PPI with NSAID are far more expensive than Arthrotec. It has been proposed that coxibs be used in acetaminophen failures[30] in OA and in some circumstances as a first-line treatment in more symptomatic OA[3,4]. We think 6 months' data are sufficient to detect cost differences between the 2 groups, as this was a 3-month-long trial followed by a "real-world" treatment for 3 months in order to allow the natural history of their OA treatment to occur after the trial. We added the costs of side effects if they required patient or physician time or the use of concomitant medications. In a group at high risk for GI events, there should still be very few clinically diagnosed ulcers or complications, as our N of 1 group was given a cytoprotective agent in their intervention, and perhaps in conventional treatment, if traditional NSAID were given with PPI, the costs would have been higher. Increased side effects in the N of 1 group may have been due to reporting bias and differential dropout. More subsets in the conventional treatment group missed scheduled visits. Subjects in the N of 1 group were asked meticulously about side effects, thus they may have overreported them.

We conclude that there were no significant between-group differences in efficacy within this small trial. The cost of N of 1 trials are intrinsically higher due to increased cost of formulating a placebo, the time spent by the nurse and the physician, and the time spent by the subject completing efficacy questionnaires or other outcome measurements. They may yield better improvements than standard care, which in this case were not statistically significant. NSAID are often effective in OA (even in those who are uncertain of their efficacy), and thus N of 1 trials of NSAID did not appear to be cost effective in our selected patients with OA.

REFERENCES

1. Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. Part I. Osteoarthritis of the hip. American College of Rheumatology. Arthritis Rheum 1995;38:1535-40.
2. Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. Arthritis Rheum 1995;38:1541-6.
3. Tannenbaum H, Peloso PMJ, Russell AS, Marlow B. An evidence-based approach to prescribing NSAIDs in the treatment of osteoarthritis and rheumatoid arthritis: The Second Canadian Consensus Conference. Can J Clin Pharmacol 2000;7 Suppl A:4A-16A.
4. American College of Rheumatology Subcommittee on Osteoarthritis Guidelines. Recommendations for the medical management of osteoarthritis of the hip and knee. Arthritis Rheum 2000;43:1905-15.
5. Prashker MJ, Cahill LA, Meenan RF. Rheumatology costs are increasing faster than inflation: results of a rheumatology cost of treatment index (RCTI) [abstract]. Arthritis Rheum 1993;36 Suppl:S65.
6. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy — randomized trials in individual patients. New Eng J Med 1986;314:889-92.
7. Guyatt D, Sackett D, Adachi J, Roberts R, Chong J, Rosenbloom D, Keller J. A clinician's guide for conducting randomized trials in individual patients. Can Med Assoc J 1988;139:498-503.
8. Keller JL, Guyatt GH, Roberts RS, Adachi JD, Rosenbloom D. An N of 1 service: applying the scientific method in clinical practice. Scand J Gastroenterol 1988;147 Suppl:22-9.
9. Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MR. The N of 1 randomized controlled trial: clinical usefulness. Ann Intern Med 1990;112:293-9.
10. McLeod RS, Taylor DW, Cohen Z, Cullen JB. Single-patient randomised clinical trial. Use in determining optimum treatment for patient with inflammation of Kock continent ileostomy reservoir. Lancet 1986;1:726-8.
11. Robin ED, Burke CM. Single-patient randomized clinical trial: opiates for intractable dyspnea. Chest 1986;90:888-92.
12. Woolf GM, Townsend M, Guyatt G. Treatment of cryptosporidiosis with spiramycin in AIDS. An "N of 1" trial. J Clin Gastroenterol 1987;9:632-4.
13. Hoogwerf BF. Amitriptyline treatment of painful diabetic neuropathy: an inadvertent single-patient clinical trial [letter]. Diabetes Care 1985;8:526-7.
14. Jaeschke R, Adachi J, Guyatt G, Keller J, Wong B. Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N of 1 randomized controlled trials. J Rheumatol 1991;18:447-51.
15. Patel A, Jaeschke R, Guyatt G, Keller JL, Newhouse MT. Clinical usefulness of N-of-1 randomized controlled trials in patients with nonreversible chronic airflow limitation. Am Rev Respir Dis 1991;144:962-4.
16. Larson EB, Ellsworth AJ, Oas J. Randomized trials in single patients during a two-year period. JAMA 1993;270:2708-12.
17. March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. N of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. BMJ 1994;309:1041-5.
18. Pincus T, Swearingen C, Cummins P, Callahan LF. Preference for nonsteroidal antiinflammatory drugs versus acetaminophen and concomitant use of both types of drugs in patients with osteoarthritis. J Rheumatol 2000;27:1020-7.
19. Wolfe F, Zhao S, Lane N. Preference for nonsteroidal antiinflammatory drugs over acetaminophen by rheumatic disease patients: a survey of 1,799 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Arthritis Rheum 2000;43:378-85.
20. Lanes SF, Lanza LL, Radensky PW, et al. Resource utilization and cost of care for rheumatoid arthritis and osteoarthritis in a managed care setting: the importance of drug and surgery costs. Arthritis Rheum 1997;40:1475-81.
21. Fries JF, Spitz P, Kraines G, et al. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
22. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt L. Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes following total hip or knee arthroplasty in osteoarthritis. J Rheumatol 1988;14:1108-14.
23. Drummond M, O'Brien B. Clinical importance, statistical significance and the assessment of economic and quality of life outcomes. Health Econ 1993;2:205-12.
24. Lesquesne MG. The algofunctional indices for hip and knee osteoarthritis. J Rheumatol 1997;24:779-81.
25. Jenkinson C, Coulter A, Wright L. Short Form 36 (SF-36) health survey questionnaire: Normative data for adults of working age.

BMJ 1993;306:1437-40.
26. Scott-Lennox JA, McLaughlin-Miley C, Lennox RD, et al. Stratification of flare intensity identifies placebo responders in a treatment efficacy trial of patients with osteoarthritis. Arthritis Rheum 2001;44:1599-607.
27. Pincus T, Koch GG, Sokka T, et al. A randomized, double-blind, crossover clinical trial of diclofenac plus misoprostol versus acetaminophen in patients with osteoarthritis of the hip and knee. Arthritis Rheum 2001;44:1587-98.
28. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology Oxford 1999;38:870-7.
29. Bellamy N, Campbell J, Haraoui B, et al. Development of the Australian/Canadian (auscan) Osteoarthritis (OA) Hand Index [abstract]. Arthritis Rheum 1997;40 Suppl:S110.
30. Schnitzer TJ. Osteoarthritis management: the role of cyclooxygenase-2 selective inhibitors. Clin Ther 2001;23:313-26.