J Clin Epidemiol Vol. 52, No. 2, pp. 113–122, 1999

0895-4356/99/ $–see front matter
PII S0895-4356(98)00149-8

Are Randomized Control Trial Outcomes Influenced by the Inclusion of a Placebo Group?: A Systematic Review of Nonsteroidal Antiinflammatory Drug Trials for Arthritis Treatment

Paula A. Rochon,[1,] Malcolm A. Binns,[2] Jason A. Litner,[3] Geoffrey M. Litner,[3] Michelle S. Fischbach,[3] David Eisenberg,[4,5] Ted J. Kaptchuk,[4] William B. Stason,[6] and Thomas C. Chalmers[6] (deceased)*

[1]Kunin-Lunenfeld Applied Research Unit, Baycrest Centre for Geriatric Care, Institute for Clinical Evaluative Sciences, Departments of Medicine and Public Health Sciences, Clinical Epidemiology and Health Care Research Program, University of Toronto, Toronto, Ontario, Canada; [2]Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada; [3]Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada; [4]Center for Alternative Medicine Research, Beth Israel Hospital, Boston, Massachusetts; [5]Harvard Medical School, Boston, Massachusetts; and [6]MetaWorks, Boston, Massachusetts

ABSTRACT. Placebo groups are often included in randomized control trials evaluating drug therapy, yet we know little about the placebo effect. The purpose of our study was to evaluate how the presence of a placebo group in a randomized control trial (RCT) influences the patients' ratings of the efficacy of an active drug therapy and their reporting of its adverse effects. We identified studies published between 1966 and 1994 using MEDLINE. Randomized control trials evaluating acetylsalicylic acid, diclofenac, or indomethacin for the treatment of osteo or rheumatoid arthritis were included in our sample. Two investigators independently extracted data. Fifty-eight treatment arms met our inclusion criteria and were available for analysis. Twenty-five treatment arms evaluated a nonsteroidal antiinflammatory drug (NSAID) in placebo control trials and 33 in comparative trials. Using a logistic regression model to adjust for the differences between the evaluated drugs and between the types of arthritis, we found that patients receiving an NSAID in a placebo control trial were more likely to withdraw due to inefficacy (OR = 1.3; 95% CI, 1.0 to 1.6; P = 0.04). Using a similar model, withdrawals due to adverse effects were found to be more common when the NSAID was given in trials that did not include a placebo group (OR = 1.5; 95% CI, 1.1 to 1.9; P = 0.002) as were reports of cutaneous (OR = 4.2; 95% CI, 1.7 to 9.9), gastrointestinal (OR = 1.6; 95% CI, 1.3 to 2.0), and other types (OR = 5.3; 95% CI, 3.8 to 7.4) of adverse effects. Although reports of central nervous system adverse effects were more frequent in the comparative trials, this difference was not significant. Including a placebo group in a RCT changes how patients rate the efficacy and adverse effects of their therapy. Our results highlight the need to consider the placebo effect in the design and analyses of clinical trials. J CLIN EPIDEMIOL 52;2:113–122, 1999. © 1999 Elsevier Science Inc.

KEY WORDS. Randomized control trials, placebo therapy, arthritis, NSAIDs, systematic review, study design

INTRODUCTION

Randomized control trials (RCTs) are considered the "gold standard" for evaluating a drug therapy. In the United States, the Food and Drug Administration often requires the inclusion of a placebo control group in RCTs designed to approve new drugs for the treatment of conditions of moderate severity such as arthritis [1]. Despite obvious interest in how placebos work in clinical trials, scientific understanding of the placebo has changed little over the past 50 years.

A series of recent review articles illustrates our ongoing intrigue with the placebo. Topics of discussion include the difficulty in defining a placebo [2], how to use the placebo in the health care setting [3], speculation on how placebos may work in a variety of settings that include clinical trials [4–7], and the role of the placebo in complementary medicine [8], general practice [9], psychiatry [10], and surgery [11]. Relatively few original investigations, however, have evaluated the factors that determine how patients with medical conditions respond to a placebo therapy [12–15],

*Address correspondence to: Paula A. Rochon, MD, MPH, FRCP(C), Baycrest Centre for Geriatric Care, 3560 Bathurst Street, Toronto, Ontario, Canada, M6A 2E1.

Accepted for publication on 13 October 1998.

the role that physicians' expectations play in making a patient believe that they are receiving an effective therapy [16,17], or how study design factors may influence the effectiveness of a placebo [18–27].

We wanted to evaluate factors related to the study design that may influence outcomes of a RCT. Of particular interest was the influence of including a placebo therapy in a clinical trial on patient expectations of the effectiveness of active therapy and their reporting of adverse drug effects. We asked whether patients evaluate the efficacy and toxicity of a drug therapy differently in the setting of a placebo control trial as opposed to the same drug therapy administered in a comparative drug trial. We hypothesized that in a placebo control trial, a drug therapy would be rated as "less efficacious" and "less toxic" because the patients would known that there is up to a 50% chance that they are receiving an inert therapy, which should not improve symptoms or have adverse effects. Conversely, we hypothesized that patients participating in a comparative drug trial would know that they are receiving some form of an active therapy and thus rate the therapy as being both more effective and more toxic.

Meaningful study of the placebo effect requires that the symptoms of the condition being evaluated be of variable intensity [28]. We therefore chose to study patients with arthritis because of the fluctuating and subjective nature of the symptoms, particularly pain, associated with this condition [15]. To answer these questions, we evaluated published trials to determine the difference in the reporting of efficacy and adverse effects by patients with arthritis receiving the same nonsteroidal antiinflammatory drug (NSAID) either in a placebo or a comparative trial.

METHODS

Study Identification and Selection

To obtain our sample, we conducted a systematic review using a previously described MEDLINE search strategy to identify NSAID trials published between 1966 and May 1990 [29,30]. To identify articles published between 1990 and the end of 1994, the following MEDLINE search strategy was used: explode *antiinflammatory agents, nonsteroidal* or *NSAID* text word; limit to human, English, and adult; limit to clinical trial or randomized controlled trial; explode research design; explode *osteoarthritis*, explode *arthritis rheumatoid*.

We selected only those trials in which patients with osteo or rheumatoid arthritis were randomly allocated (word random in text) to receive an NSAID. Studies were excluded if they were meta-analyses, if the patients did not receive an oral preparation of an NSAID, if the duration of therapy was 4 days or less, if efficacy outcomes or toxicity measures were not the major focus of the article (e.g., studies evaluating cytoprotection with misoprostol), if the study represented a duplicate publication, or if the patients were not adults.

We divided the trials into two groups according to whether the NSAID was evaluated against a placebo arm or not as outlines schematically in Figure 1. Of the 41 trials in the NSAID versus Placebo group, we identified those that included the three most frequently evaluated NSAIDs and provided efficacy information for further evaluation. These NSAIDs were acetylsalicyclic acid ($n = 8$ trials), diclofenac ($n = 6$ trials), and indomethacin ($n = 9$ trials). Some trials analyzed more than one of the three identified drugs or studied different formulations of the same drug. We therefore used treatment arms rather than trials as the unit of observation. The NSAID versus Placebo sample consisted of 25 treatment arms (Table 1). Similarly, for the NSAID versus Comparative Drug group, we identified the subset of trials that compared either acetylsalicyclic acid, diclofenac, or indomethacin to another drug but not to a placebo. The NSAID versus Comparative Drug sample consisted of 33 treatment arms (Table 2). The raw data used for the analyses are included in Appendix 1.

Data Abstraction

Data on each trial were collected in duplicate by two reviewers to ensure accuracy. Discrepancies in data collection were resolved in a consensus meeting. Because an article may contain more than one trial, each trial was reviewed separately. Basic study design information collected included the type of study (placebo control or comparative drug), the drug being evaluated, the number of patients randomized in each treatment arm, the year the trial was published, the duration of the trial, and whether patients with previous NSAID-related adverse effect were ineligible for the trial. Basic patient information collected included diagnosis (osteo or rheumatoid arthritis), the gender distribution, and mean age.

Efficacy

Efficacy was assessed by measuring the number of patients who withdrew from each trial for the sole reason that their treatment was not working. Studies that did not specify the number of patients who were randomized to each treatment group or the number of patients who withdrew due to inefficacy were excluded from this assessment.

Adverse Effects

Adverse effects were assessed by two methods. First, we recorded the total number of patients who withdrew from trials as a result of an adverse drug effect. Only studies that specifically stated that patients withdrew as a result of drug toxicity were included. Second, we recorded the numbers and types of adverse events reported by patients that were

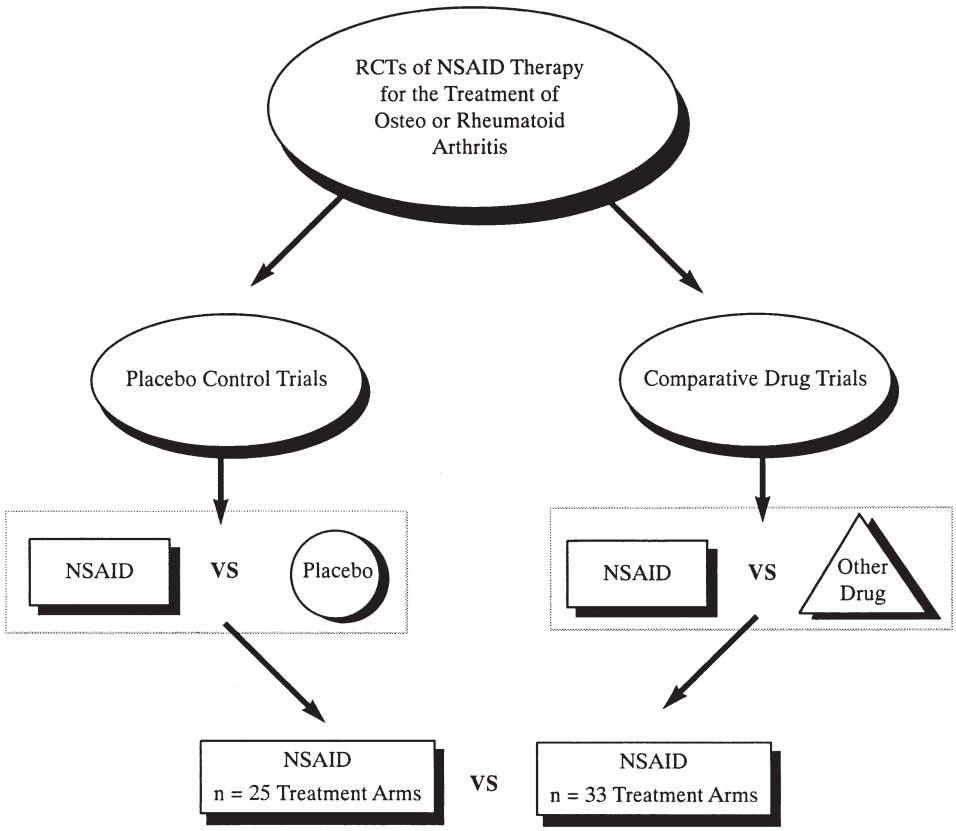

FIGURE 1. Schematic outline of the study design. This flow chart outlines the study design. All NSAID treatment arms evaluating ASA, diclofenac, or indomethacin were selected from both placebo and comparative drug trials. From placebo control trials, 25 NSAID treatment arms were available, and from comparative trials 33 treatment arms were available for comparison. The NSAID treatment arms obtained from each of these type of trials evaluated for efficacy and adverse effects.

not necessarily severe enough to lead to a withdrawal from the trial. Specifically, the number of patients reporting adverse effects in the cutaneous, central nervous system, gastrointestinal, or other systems were tabulated. Withdrawals due to adverse effects and the reporting of adverse effects by system were not necessarily mutually exclusive. For example, it may have been reported that a patient withdrew from a trial due to an adverse effect and also reported that an adverse effect in a particular system occurred. Generally, details concerning adverse effects were poorly reported.

Four treatment arms reported a total of 13 patients who withdrew for reasons of both lack of efficacy and adverse effects. These patients were not included in our analyses.

TABLE 1. NSAID versus placebo sample selection

Description	*n*
Osteo or rheumatoid arthritis	577
Exclusions	
No placebo group in parallel design	450
No NSAID comparison group	34
Other (nonrandom allocation [word random required in text] [12], nonoral therapy [4], efficacy outcome not major focus [24], duplicate publications [7], nonadult population [1])	48
Not drug of interest	22
Trials remaining	23
Treatment arms[a]	28
Treatment arms remaining[b]	25

[a]Treatment arms were used as the unit of observation because some trials analyzed more than one of the three drugs of interest or different formulations of the same drug.

[b]Three treatment arms were excluded because they did not provide the necessary data (numbers of patients randomized or reports of withdrawals/adverse effects) required to calculate the odds ratio.

TABLE 2. NSAID versus comparative drug sample selection

Description	*n*
Osteo or rheumatoid arthritis	577
Exclusions	
Not a comparative trial with parallel design	262
Not drug of interest	243
Other (nonrandom allocation [word random required in text] [9], meta-analysis [7], nonoral therapy [2], duration of therapy 4 days or less [1], efficacy outcome not major focus [16], duplicate publications [2], nonadult population [1]	38
Trials remaining	34
Treatment arms[a]	38
Treatment arms remaining[b]	33

[a]Treatment arms were used as the unit of observation because some trials analyzed more than one of the three drugs of interest or different formulations of the same drug.

[b]Five treatment arms were excluded because they did not provide the necessary data (numbers of patients randomized or reports of withdrawals/adverse effects) required to calculate the odds ratio.

Statistical Analyses

Differences in continuous variables between NSAID treatment arms in a placebo setting and those in a comparative drug setting were evaluated using both weighted (by sample size) and nonweighted *t* tests. We used weighted (by sample size) and unweighted Pearson χ^2 tests to examine differences in categorical variables between NSAID treatment arms in a placebo setting and those in a comparative drug setting. The Wilcoxon rank sum statistic was used to evaluate the difference between the median years of publication of these two groups. Calculations were performed with SAS [31].

Multiple logistic regression analysis was used, adjusting for differences in characteristics of the treatment arms and type of patients participating, to evaluate the independent effect of the study design (i.e., placebo versus comparative drug setting) on withdrawals due to inefficacy. The following variables were included in the models: type of study design (placebo or comparative), the NSAID evaluated (acetylsalicylic acid, diclofenac, or indomethacin), and the type of arthritis being studied (osteo or rheumatoid arthritis). The year that the trial was conducted was not included in the logistic regression model because this variable was related to our independent measure of study design. Placebo trials were more likely to be published in the more distant past. This multiple logistic regression was conducted using the SAS system, specifically the logistic procedure with the "events/trials" syntax [31].

To evaluate the independent effect of the study design on withdrawals due to adverse effects and the reporting of adverse effects in each of four systems (cutaneous, central nervous system, gastrointestinal, or other systems), identical logistic regression models were used for each of the independent variables.

We considered P values less than 0.05 to indicate statistically significant findings.

RESULTS

Characteristics of the trials and the types of patients being studied in the NSAID versus Placebo and the NSAID versus Comparative Drug groups are outlined in Table 3. Treatment arms evaluated in the NSAID versus Placebo sample tended to be less likely to evaluate diclofenac (χ^2_2 = 2.95; P = 0.23) and were published earlier (Wilcoxon S = 474.5; P = 0.0001) when compared with the treatment arms used in the NSAID versus Comparative Drug sample. No difference was found in the average duration of the trials. Exclusion criteria used in comparative trials tended to be more restrictive. Only 12 (48%) NSAID versus Placebo treatment arms listed a previous adverse effect attributed to an NSAID as an exclusion criteria compared with 22 (67%) NSAID versus Comparative Drug treatment arms (χ^2_1 = 2.0; P = 0.15). There was a trend for treatment arms in the NSAID versus Placebo sample to study rheumatoid arthritis (χ^2_1 = 2.92; P = 0.09). The distribution of males and females and the average age of participants were similar between the groups.

Efficacy Data

Twenty-five treatment arms representing an average of 58 patients each were evaluated in the NSAID versus Placebo setting. Thirty-three treatment arms representing an average of 75 patients each were evaluated in the NSAID versus Comparative Drug setting (Table 3).

Of those treatment arms specifying the number of patients who withdrew due to inefficacy, 144 (10.3%) of 1396 patients receiving active therapy in the NSAID versus Placebo group withdrew due to inefficacy compared with 182 (9.8%) of 1850 patients receiving an active therapy in the NSAID versus Comparative Drug sample.

We used a logistic regression model to adjust for the differences between groups with regards to the types of drugs used and whether the patients were diagnosed with osteo or rheumatoid arthritis (Table 4). Patients participating in the NSAID versus Placebo group were more likely to withdraw because they claimed the drug was not effective compared with patients participating in NSAID versus Comparative Drug trials (OR = 1.3; 95% confidence interval [CI], 1.0 to 1.6; P = 0.04).

Of the three drugs evaluated, indomethacin was found to be more effective than diclofenac or acetylsalicyclic acid. Using indomethacin as the reference drug, patients were significantly more likely to leave trials because they considered the drug to be ineffective when they received either acetylsalicylic acid (OR = 2.3; 95% CI, 1.7 to 3.3, P = 0.0001) or diclofenac (OR = 2.2; 95% CI, 1.6 to 3.1; P = 0.0001).

Adverse Events

One hundred fourteen (9.3%) of 1232 patients receiving active therapy in the NSAID versus Placebo setting reported an adverse drug effect serious enough to cause them to withdraw from their trial compared with 242 (13.0%) of 1862 patients who received an NSAID in the NSAID versus Comparative Drug setting.

Using a logistic regression model to adjust for differences in treatment arm characteristics, we evaluated withdrawals due to adverse effects and adjusted for the differences between the placebo and comparative groups in regards to the types of NSAID evaluated and whether the patients were diagnosed with osteo or rheumatoid arthritis (Table 5). We found that withdrawals due to adverse effects were greater when the NSAID was being given in the comparative drug setting (OR = 1.5; 95% CI, 1.1 to 1.9; P = 0.002). Patients receiving indomethacin (OR = 1.7; 95% CI, 1.3 to 2.2; P = 0.0003) or acetylsalicylic acid (OR = 2.7; 95% CI, 1.9 to

TABLE 3. Characteristics of the treatment arms in the two study design groups

	Group				
	NSAID versus Placebo	NSAID versus Comparative	P value	Weighted P value[a]	Test
NSAID elevaluated by Tx arm[b,c]					
ASA	6 (24%)	8 (24%)	0.23	0.001	χ^2
Diclofenac	8 (32%)	17 (52%)			
Indomethacin	11 (44%)	8 (24%)			
Number of Patients per Tx arm[b]					
Average	57.9	75.4	0.22	N/A	*t* test
Arthritis type studied in Tx arm[b,c]					
Osteo	8 (32%)	18 (55%)	0.09	0.22	χ^2
Rheumatoid	17 (68%)	15 (45%)			
Year of publication					
Median	1983	1989	0.0001	N/A	Wilcoxon rank sum test
Duration of trial (weeks)					
Average	9.7	9.9	0.96	0.47	*t* test
Exclusion due to previous NSAID adverse effect	12 (48%)	22 (67%)	0.15	N/A	*t* test
Proportion of males in Tx arm[b]					
Average	0.26	0.29	0.58	0.5	*t* test
Mean age of subjects in Tx arm[b]					
Average	56.9	58.4	0.42	0.28	*t* test

[a]Weighted by the number of subjects in each treatment arm.
[b]Treatment arms were used as the unit of observation because some trials analyzed more than one of the three drugs of interest or different formulations of the same drug.
[c]Numbers are counts of treatment arms with proportion of group in parentheses.
Abbreviation: NSAID; nonsteroidal antiinflammatory drug.

3.7; P = 0.0001) were more likely than those receiving diclofenac to leave a trial as a result of an adverse effect.

Reporting of adverse effects that were not necessarily severe enough to lead to withdrawal followed a similar pattern. Percentages of patients who reported an adverse effect in each of the four systems evaluated (central nervous system, cutaneous, gastrointestinal, or other) by type of study design and NSAID are outlined in Figure 2a, b, c, and d. Using our regression model as outlined in Table 6, we found that cutaneous (OR = 4.2; 95% CI, 1.7 to 9.9; P = 0.001), gastrointestinal (OR = 1.6; 95% CI, 1.3 to 2.0; P = 0.0001), and other adverse effects (OR = 5.3; 95% CI, 3.8 to 7.4; P =

TABLE 4. Results of multiple logistic regression model to evaluate the effect of the study design, controlling for the NSAID evaluated and type of arthritis, on withdrawals due to inefficacy

Variable	Odds ratio	95% CI	P value
Study design			
Placebo control trial	1.3	1.0 to 1.6	0.04
Comparative drug trial	1[a]		
NSAID evaluated			
ASA	2.3	1.7 to 3.3	0.0001
Diclofenac	2.2	1.6 to 3.1	0.0001
Indomethacin	1[a]		
Arthritis type			
Osteoarthritis	0.20	0.14 to 0.29	0.0001
Rheumatoid	1[a]		

[a]Reference category for calculating odds ratio.
Abbreviations: CI, Confidence interval; NSAID, nonsteroidal antiinflammatory drug; ASA, acetylsalicylic acid.

TABLE 5. Results of multiple logistic regression model to evaluate the effect of the study design, controlling for the NSAID evaluated and type of arthritis, on withdrawals due to adverse effects

Variable	Odds ratio	95% CI	P value
Study design			
Comparative drug trial	1.5	1.1 to 1.9	0.002
Placebo control trial	1[a]		
NSAID evaluated			
ASA	2.7	1.9 to 3.7	0.0001
Indomethacin	1.7	1.3 to 2.2	0.0003
Diclofenac	1[a]		
Arthritis type			
Osteoarthritis	1.5	1.1 to 1.9	0.003
Rheumatoid	1[a]		

[a]Reference category for calculating odds ratio.
Abbreviations: CI, confidence interval; NSAID, nonsteroidal antiinflammatory drug; ASA, acetylsalicylic acid.

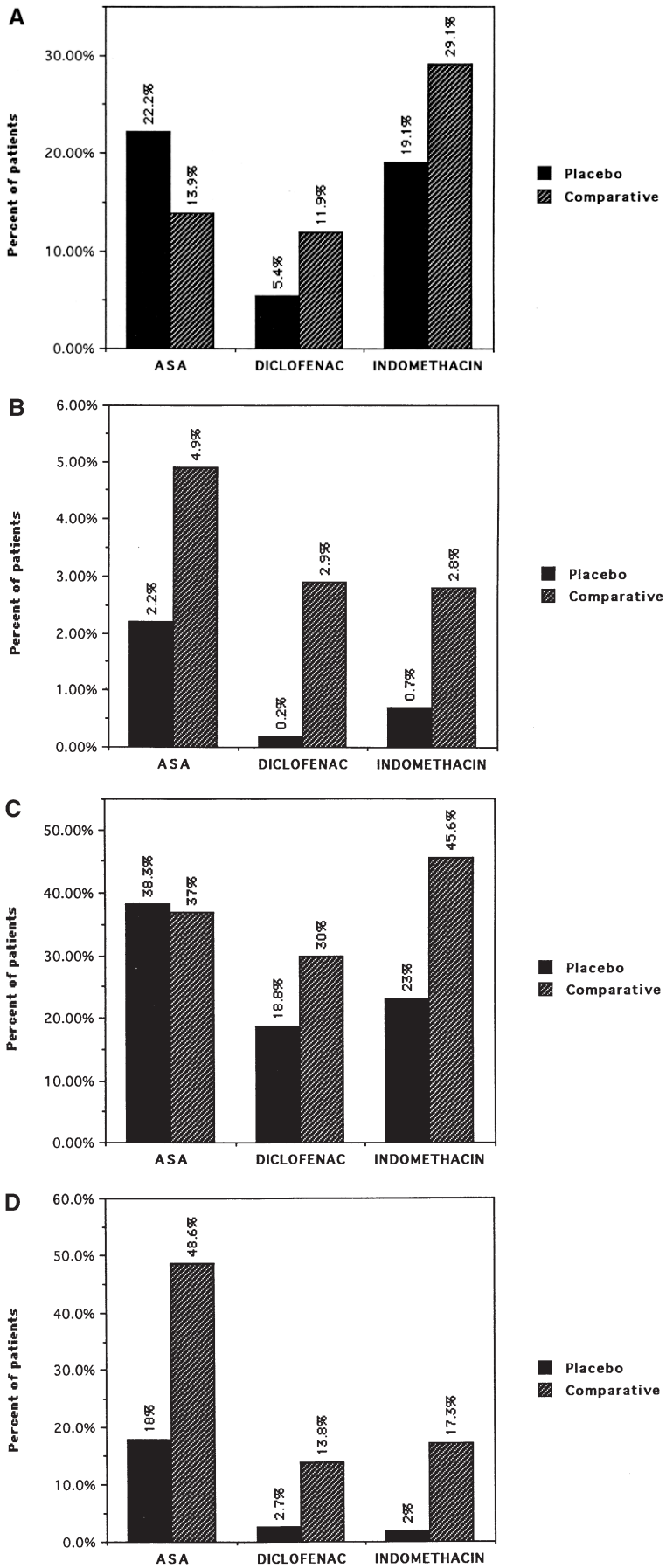

FIGURE 2. (a) Central nervous system adverse effects. The percentage of patients reporting a central nervous system adverse effect when either acetylsalicylic acid, diclofenac or indomethacin were administered in the setting of a placebo or a comparative trial. (b) Cutaneous adverse effects. The percentage of patients reporting a cutaneous adverse effect when either acetylsalicyclic acid, diclofenac or indomethacin were administered in the setting of a placebo or a comparative trial. (c) Gastrointestinal adverse effects. The percentage of patients reporting a gastrointestinal adverse effect when either acetylsalicylic acid, diclofenac or indomethacin were administered in the setting of a placebo or a comparative trial. (d) Other adverse effects. The percentage of patients reporting other types of adverse effect when either acetylsalicylic acid, diclofenac or indomethacin were administered in the setting of a placebo or a comparative trial. *Note*: A patient may have reported multiple adverse effects; this could not be accounted for in our study.

TABLE 6. Results of four multiple logistic regression models to evaluate the effect of the study design, controlling for the NSAID evaluated and type of arthritis, on the reporting of adverse effects by body systems

Variable	Odds ratio	95% CI	P value
Central nervous system			
Study design			
Comparative drug trial	1.3	0.98 to 1.6	0.08
Placebo control trial	1[a]		
NSAID evaluated			
ASA	0.49	0.36 to 0.66	0.0001
Diclofenac	0.32	0.24 to 0.43	0.0001
Indomethacin	1[a]		
Arthritis type			
Osteoarthritis	0.56	0.42 to 0.75	0.0001
Rheumatoid	1[a]		
Cutaneous			
Study design			
Comparative drug trial	4.2	1.7 to 9.9	0.001
Placebo control trial	1[a]		
NSAID evaluated			
ASA	1.5	0.72 to 3.1	0.28
Diclofenac	1.1	0.51 to 2.4	0.78
Indomethacin	1[a]		
Arthritis type			
Osteoarthritis	0.27	0.11 to 0.68	0.005
Rheumatoid	1[a]		
Gastrointestinal			
Study design			
Comparative drug trial	1.6	1.3 to 2.0	0.0001
Placebo control trial	1[a]		
NSAID evaluated			
ASA	1.0	0.81 to 1.3	0.79
Diclofenac	0.53	0.43 to 0.66	0.0001
Indomethacin	1[a]		
Arthritis type			
Osteoarthritis	1.2	1.0 to 1.5	0.06
Rheumatoid	1[a]		
Other			
Study design			
Comparative drug trial	5.3	3.8 to 7.4	0.0001
Placebo control trial	1[a]		
NSAID evaluated			
Diclofenac	0.12	0.083 to 0.17	0.0001
Indomethacin	0.18	0.13 to 0.26	0.0001
ASA	1[a]		
Arthritis type			
Osteoarthritis	1.6	1.1 to 2.2	0.009
Rheumatoid	1[a]		

[a]Reference category for calculating odds ratio.

Abbreviations: CI, confidence interval; NSAID, nonsteroidal anti-inflammatory drug; ASA, acetylsalicylic acid.

0.0001) were significantly more likely to be reported by patients receiving an active therapy in a comparative trial setting compared to patients who received the same active therapy in the setting of a placebo trial. While the reports of central nervous system adverse effects were more frequent in the comparative trials, this difference was not significant (OR = 1.3; 95% CI, 0.98 to 1.6; P = 0.08).

DISCUSSION

Efficacy

We demonstrate that the study design, that is whether or not a patient is participating in a study that includes a placebo group, may influence the outcome of an RCT. Patients receiving the same NSAID are less likely to rate their therapy as being efficacious when it is administered in the setting of a placebo trial compared with when the same drug is administered in a comparative drug setting.

Gotzsche [18] conducted a similar meta-analysis of NSAID trials to evaluate the influence of including a placebo group in a clinical trial on the patient's expectations of the effectiveness of the active therapy. In comparison to our study, Gotzsche included only trials published up to 1988, evaluated patients with rheumatoid arthritis, and used the tender joint counts as the outcome measure. Gotzsche found that the number of tender joints decreased by a slightly greater amount in patients receiving an NSAID in a comparative trial than when patients received the same NSAID in a placebo trial. The small difference in joint count reported by Gotzsche between those receiving active treatment in the placebo and the comparative trial setting was not statistically significant.

Adverse Effects

We demonstrate that the study design, in this case whether or not the patient receives an NSAID as part of a placebo or a comparative trial, may impact on the number of patients who withdraw from the study due to an adverse effect. Our results suggest that patient withdrawals due to an adverse effect were less likely to occur if the active treatment was being administered in a placebo control trial. In this setting, patients may have been more reluctant to consider their adverse effect substantial, presumably because they know that they may be receiving an inert therapy. Conversely, patients participating in a comparative drug trial, in which we are assuming they know they were receiving some form of active therapy, were more likely to withdraw because they thought their therapy might be toxic.

Likewise, when the reporting of adverse effects that were not necessarily severe enough to lead to study withdrawal were evaluated, patients receiving an active therapy in a placebo control trial reported fewer adverse effects for each body system. This suggests that patients' reports of adverse effects are influenced by their knowledge of the likelihood of receiving a placebo therapy.

Limitations

The results we have obtained likely underestimate the influence that study design has on patient ratings of their drug treatment. Because we evaluated a large number of trials that were not designed to measure how study design may influence trial outcomes, we could not control for a number

of factors that may be important. For example, standard information was not reported on patients' understanding of their participation in a placebo control versus a comparative trial. Bergmann *et al.* [26] found that when patients gave their informed consent the analgesic effect of medications changed. They report that informed consent increased the efficacy of both naproxen and placebo and decreased the difference between the efficacy of these two medications. Owing to the work of ethics and research review boards, informed consent is now required from study participants. Therefore patients now participating in clinical trials will better understand the types of treatment they may receive.

Our findings may underestimate the true difference in adverse effects experienced in the placebo and comparative drug trials. The trials we evaluated were not designed to measure the influence of study design on the type of patient selected for inclusion in the trial. For example, we could measure whether our comparative drug sample contained patients who may be less likely to develop an NSAID related adverse effect. While not significant, more comparative trials in our study excluded patients who had previously experienced an NSAID-related adverse drug effect. This suggests that investigators may be more cautious about including patients in a comparative drug trial, in which they known that the patient will receive some form of active therapy. There are likely many other important factors that may differ between placebo control and comparative trials that we have not measured.

In general, the reporting of trial outcomes in our sample was inconsistent. For example, not all trials reported withdrawals due to inefficacy or adverse effects. Of those trials that did report withdrawals, some reported withdrawals due to inefficacy, others reported withdrawals due to adverse effects, and still others reported a category that combined these two reasons. Previous work [32] has shown that the quality of randomized control trials evaluating NSAIDs for the treatment of arthritis, as reflected by quality scores, is relatively low. Our findings highlight the need for standard reporting of outcomes in randomized control trials [33].

Practical Applications

Randomized control trials are designed to decrease the likelihood of making the mistake of calling a therapy efficacious when it is not, or calling a therapy safe when it is not. Choosing an appropriate study design for clinical trials is key, and based on our research findings, one that should be determined by the major outcome of interest for the investigator. Our findings suggest that even when a randomized placebo control trial is feasible, it may not always be the best study design for the evaluation of both efficacy and adverse effects. If efficacy is the outcome of interest, placebo control trials may be the most appropriate study design because patients participating in placebo drug trials are more conservative in rating a drug as being effective. If adverse effects are the major outcome of interest, comparative trials may be the preferred design because patients participating in a comparative drug trial are more liberal in reporting adverse drug effects.

Additionally, our results highlight the need to systematically identify and evaluate the factors related to the design of placebo control trials that may influence the final outcome. Once these factors that may change the way patients report the efficacy and adverse effects of medications are understood, standard methods should be developed to control for them in the design of future placebo control clinical trials [19].

CONCLUSION

The beliefs and expectations of patients are among the more important factors responsible for the placebo effect [16,34]. We demonstrate how a change to the study design, including or not including a placebo group, may influence patients' beliefs and expectations about the probable efficacy and the toxicity of the drug they are receiving. Our results suggest that the placebo effect may change the way patients rate the performance of the drug in the trial and support the need to revisit the way that placebos affect the outcomes of clinical trials.

We are forever grateful to the late Dr. T.C. Chalmers who inspired this project and to Mrs. T.C. Chalmers who so graciously facilitated this work. The authors thank Iain Chalmers, Peter Gotzsche, Donald Stuss, and Michael Gordon for their comments on earlier drafts of the manuscript; Christopher H. Schmid for advice regarding statistics; and Jocalyn Clark, Penelope de Nobrega, and Bruce Kupelnick for their assistance with this project.

References

1. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. **N Engl J Med** 1994; 331: 394–398.
2. Gotzsche PC. Is there logic in the placebo? **Lancet** 1994; 344: 925–926.
3. Chaput de Saintonge DM, Herxheimer A. Harnessing placebo effects in health care. **Lancet** 1994; 344: 995–998.
4. Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. **JAMA** 1994; 271: 1609–1614.
5. Kleijnen J, de Craen AJM, van Everdingen J, Krol L. Placebo effect in doubleblind clinical trials: A review of interactions with medications. **Lancet** 1994; 344: 1347–1349.
6. Oh VMS. Magic or medicine? Clinical pharmacological basis of placebo medication. **Ann Acad Med Singapore** 1991; 20: 31–37.
7. Weiner M, Weiner GJ. The kinetics and dynamics of responses to placebo. **Clin Pharm Ther** 1996; 60: 247–254.
8. Joyce CRB. Placebo and complementary medicine. **Lancet** 1994; 344: 1279–1281.
9. Thomas KB. The placebo in general practice. **Lancet** 1994; 344: 1066–1067.
10. Laporte JR, Figueras A. Placebo effects in psychiatry. **Lancet** 1994; 344: 1206–1209.

11. Johnson AG. Surgery as a placebo. **Lancet** 1994; 344: 1140–1142.
12. Langley GB, Sheppeard H, Wigley RD. Placebo therapy in rheumatoid arthritis. **Clin Exp Rheumatol** 1983; 1: 17–21.
13. Lasagna L, Mosteller F, von Felsinger JM, Beecher HK. A study of the placebo response. **Am J Med** 1954; 16: 770–779.
14. McQuay H, Carroll D, Moore A. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. **Pain** 1995; 64: 331–335.
15. Traut EF, Passarelli EW. Placebos in the treatment of rheumatoid arthritis and other rheumatic conditions. **Ann Rheum Dis** 1957; 16: 18–21.
16. Gracely RH, Dubner R, Deeter WR, Wolskee PJ. Clinicians' expectations influence placebo analgesia. **Lancet** 1985; 1: 43.
17. Affleck DC, Eaton MT, Mansfield E. The action of a medication and the physician's expectations. **Nebr State Med J** 1966; 51: 331–334.
18. Gotzsche PC. Meta-analysis of NSAIDs: Contribution of drugs, doses, trial designs, and meta-analytic techniques. **Scand J Rheum** 1993; 22: 255–260.
19. Rosenzweig P, Brohier S, Zipfel A. The placebo effect in healthy volunteers: Influence of experimental conditions on the adverse events profile during phase I studies. **Clin Pharmacol Ther** 1993; 54: 578–583.
20. Skovlund E. Should we tell trial patients that they might receive placebo? **Lancet** 1991; 337: 1041.
21. Bystritsky A, Waikar SV. Inert placebo versus active medication: Patient blindability in clinical pharmacological trials. **J Nerv Ment Dis** 1994; 182: 485–487.
22. Tangrea JA, Adrianza ME, Helsel WE. Risk factors for the development of placebo adverse reactions in a multicenter clinical trial. **Ann Epidemiol** 1994; 4: 327–331.
23. Levine J, Gordon N. Influence of the method of drug administration on analgesic response. **Nature** 1984; 312: 755–756.
24. Moscucci M, Byrne L, Weintraub M, Cox C. Blinding, unblinding, and the placebo effect: An analysis of patients' guesses of treatment assignment in a double-blind clinical trial. **Clin Pharmacol Ther** 1987; 41: 259–265.
25. Schapira K, McClelland HA, Griffiths NR, Newell DJ. Study on the effects of tablet colour in the treatment of anxiety states. **BMJ** 1970; 2: 446–449.
26. Bergmann JF, Chassany O, Gandiol J, Deblois P, Kanis JA, Segrestaa JM, Caulin C, Dahan R. A randomized clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. Clinical trials and meta-analysis. 1994; 29: 41–47.
27. de Craen AJM, Roos PJ, de Vries AL, Kleijnen J. Effect of color of drugs: Systematic review of perceived effect of drugs and of their effectiveness. **BMJ** 1996; 313: 1624–1626.
28. Bourne HR. Rational use of placebo. In: Melmon KL, Morelli HF, Eds. **Clinical Pharmacology: Basic principles in therapeutics, 2nd ed.** New York: Macmillan; 1978: 1052–1062.
29. Chalmers TC, Berrier J, Hewitt P, Berlin J, Reitman D, Nagalingam R, Sacks H. Meta-analysis of randomized controlled trials as a method of estimating rare complications of non-steroidal antiinflammatory drug therapy. **Ailment Pharmacol Ther** 1988; 2S: 9–26.
30. Rochon PA, Fortin PR, Dear KBG, Minaker KL, Chalmers TC. Reporting of age data in clinical trials of arthritis: Deficiencies and solutions. **Arch Intern Med** 1993; 153: 243–248.
31. SAS Institute Inc. SAS Statistical Analysis System (Release 6.11 ed.) Cary, NC: SAS Institute Inc.; 1995.
32. Rochon PA, Gurwitz JH, Simms RW, Fortin PR, Felson DT, Minaker KL, Chalmers TC. A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. **Arch Intern Med** 1994; 154: 157–163.
33. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. **JAMA** 1994; 272: 1926–1931.
34. Benson H, McCallie DP Jr. Angina pectoris and the placebo effect. **N Engl J Med** 1979; 300: 1424–1429.

APPENDIX 1. Raw data

Treatment arm[a]	Design group	Drug	Number of patients	Type of arthritis	Year	Duration (weeks)	Proportion Males	Mean Age	Withdrawals due to		Reports of Adverse Effects			
									Inefficacy	AE	CNS	GI	Cut	Other
1	pl	asa	10	oa	1983	12	0.30	60.1	3	1	4	6	0	8
2	pl	asa	105	ra	1981	8	0.32	57	19	10	13	25	4	0
3	pl	asa	65	ra	1983	12	0.28	52	12	15	32	36	—	—
4	pl	asa	9	ra	1983	12	0.00	58	0	5	4	5	0	4
5	pl	asa	7	ra	1983	12	0.14	50	1	3	0	3	0	1
6	pl	asa	52	ra	1986	6	0.44	53	—	—	2	20	0	20
7	pl	dic	68	oa	1985	2	0.00	—	0	5	1	19	0	2
8	pl	dic	208	oa	1991	4	0.45	69.2	7	10	8	28	0	4
9	pl	dic	45	oa	1993	104	0.38	63	3	6	—	—	—	—
10a	pl	dic	62	ra	1976	4	0.15	55.1	5	5	7	15	0	1
11	pl	dic	21	ra	1978	2	—	—	0	1	1	5	0	1
12	pl	dic	89	ra	1986	6	—	—	27	2	7	17	1	4
13	pl	dic	89	ra	1986	6	—	—	27	—	—	—	—	—
14a	pl	dic	75	ra	1986	10	—	—	19	—	—	—	—	—
15	pl	ind	75	oa	1977	4	0.45	53	1	10	2	13	1	1
16a	pl	ind	50	oa	1983	3	—	—	2	5	—	—	—	—
16b	pl	ind	49	oa	1983	3	—	—	2	6	—	—	—	—
16c	pl	ind	47	oa	1983	3	—	—	0	1	—	—	—	—
17	pl	ind	14	ra	1966	4	—	—	0	1	6	8	0	0
18	pl	ind	71	ra	1967	12	—	—	5	10	—	—	—	—
10b	pl	ind	63	ra	1976	4	0.21	55.5	6	8	21	14	0	2
19	pl	ind	10	ra	1980	3	—	—	2	0	—	—	—	—
20a	pl	ind	56	ra	1983	2	—	—	0	4	—	—	—	—
20b	pl	ind	55	ra	1983	2	—	—	2	4	—	—	—	—
20c	pl	ind	53	ra	1983	2	—	—	1	2	—	—	—	—
21	cp	asa	16	oa	1977	10	—	—	1	4	—	10	—	—
22	cp	asa	20	oa	1987	26	—	—	—	2	2	15	1	18
23	cp	asa	31	ra	1974	26	—	—	—	—	20	15	—	—
24a	cp	asa	47	ra	1975	2	—	—	6	13	—	—	—	—
14b	cp	asa	290	ra	1986	12	—	—	40	—	—	—	—	—
25	cp	asa	136	ra	1987	24	—	—	30	13	16	58	11	108
26	cp	asa	108	ra	1989	12	0.17	54.8	13	16	6	15	0	7
27a	cp	asa	100	ra	1989	12	0.29	53.8	—	23	11	39	6	44
28	cp	dic	13	oa	1987	4	0.08	63	0	0	0	2	0	2
29	cp	dic	46	oa	1988	8	0.33	70.7	0	12	5	28	0	0
30	cp	dic	47	oa	1989	8	0.36	61	0	3	—	—	—	—
31	cp	dic	121	oa	1990	6	—	—	—	14	10	35	—	15
32	cp	dic	40	oa	1990	12	0.28	58.3	—	7	9	18	1	—
33	cp	dic	87	oa	1992	8	0.14	59	—	4	—	—	—	—
34	cp	dic	43	oa	1992	6	0.42	57.2	2	3	1	3	0	4
35	cp	dic	32	oa	1992	4	—	—	0	3	7	4	1	3
36	cp	dic	31	oa	1992	6	0.23	57	1	2	2	10	0	3
37	cp	dic	68	oa	1993	12	0.34	61.5	0	11	—	24	—	1
38	cp	dic	62	oa	1993	12	0.19	62.8	0	13	—	25	—	—
39	cp	dic	50	oa	1993	12	0.28	62.7	2	11	18	34	2	31
14c	cp	dic	304	ra	1986	12	—	—	45	—	—	—	—	—
40	cp	dic	15	ra	1988	4	0.20	—	1	2	0	0	2	0
27c	cp	dic	111	ra	1989	12	0.35	53.3	—	3	6	18	4	9
27b	cp	dic	94	ra	1989	12	0.39	52.9	—	6	13	21	4	18
41	cp	dic	55	ra	1993	12	0.35	57.5	4	1	—	12	—	—
42	cp	ind	18	oa	1987	4	0.39	65	0	1	2	5	0	3
43	cp	ind	73	oa	1989	8	—	—	4	11	—	—	—	—
44	cp	ind	33	oa	1991	6	1.00	53.8	—	4	9	10	0	6
45	cp	ind	99	oa	1992	6	0.40	56.4	7	20	15	38	0	—
24b	cp	ind	48	ra	1975	2	—	—	4	5	—	—	—	—
46	cp	ind	93	ra	1983	8	0.18	—	5	16	25	35	4	6
47	cp	ind	70	ra	1988	12	0.34	52.6	13	4	31	19	4	11
48	cp	ind	86	ra	1991	6	0.27	55.6	4	15	34	75	3	26

[a]Some trials analyzed more than one drug or different formulations of a drug. They are given the same number but are differentiated with a letter.