

Method for Assessing Therapeutic Potential of Anti-inflammatory Antirheumatic Drugs in Rheumatoid Arthritis

PETER LEE, JOHN WEBB, JOHN ANDERSON, W. WATSON BUCHANAN

British Medical Journal, 1973, 2, 685-688

Summary

A 14-day, single-blind trial of prednisone, aspirin, and placebo was carried out in 128 patients suffering from rheumatoid arthritis, using subjective criteria only (severity of pain daily on a pain chart and assessment of the drug for effectiveness). The average treated pain rating, mean patient satisfaction rating, and mean number of days withdrawn from each drug all showed significant differences between prednisone, aspirin, and placebo. Of various pretreatment observations, only the initial pain score and articular index of joint tenderness were significantly related to the average treated pain rating.

The trial method is simple and allows many patients to participate without being time consuming for patient or physician. The method seems to have potential in comparing the comparative effectiveness of anti-inflammatory analgesics used in the treatment of patients with rheumatoid arthritis.

Introduction

The basis of current clinical trials of new antirheumatic drugs is the standard, double-blind, crossover method, the drug on trial usually being compared with a placebo or a known effective antirheumatic agent, such as aspirin, or both. The effectiveness of the drug is established on subjective and objective criteria: relief of pain experienced by the patient, joint tenderness (Ritchie *et al.*, 1968) and the number of actively inflamed joints (Lansbury, 1966), duration of morning stiffness, finger joint circumference (Hart and Clark, 1951; Boardman and Hart, 1967; Webb *et al.*, 1973), grip strength (Wright, 1959; Lee *et al.*, 1973), various composite indices (Lansbury, 1958; Mainland, 1967), thermography (Hollander *et al.*, 1951; Cosh and Ring, 1967; Collins and Cosh, 1970), and isotope "uptake" of joints (Weiss *et al.*, 1966; Whaley *et al.*, 1968; Dick *et al.*, 1970 a; Collins *et al.*, 1971), and various laboratory measurements such as the erythrocyte sedimentation rate.

The major limitation of this type of trial is that it answers only the question of whether the new antirheumatic drug is better or not than placebo or the active agent selected for comparison. Moreover, many of the clinical and laboratory indices are not relevant to the assessment of the potential therapeutic value of the drug. Indeed, it may be argued (Hart and Huskisson, 1972) that the only relevant feature of an antirheumatic drug is its ability to relieve joint pain, which is the main complaint of the patient suffering from rheumatoid arthritis, since none of the current antirheumatic drugs, including corticosteroids, sufficiently suppress inflammation to

prevent joint damage and crippling deformities (Medical Research Council and Nuffield Foundation, 1960).

The most sensitive criterion in assessing antirheumatic drugs is the patient's own assessment of pain relief (Deodhar *et al.*, 1973). We report a new method of studying antirheumatic drugs using only the patient's pain response. It is hoped that the response rating scales introduced here will go some way to setting up a standard for comparing the effects of new and established antirheumatic drugs. In the present study three preparations (placebo, prednisone, and aspirin) are compared.

Patients and Methods

Patients with rheumatoid arthritis, as defined by the diagnostic criteria of the American Rheumatism Association (Ropes *et al.*, 1959), were invited to participate in the study. All of these patients had previously attended as outpatients or had received treatment as inpatients at the Centre for Rheumatic Diseases. Patients who were currently receiving corticosteroid or corticotrophin therapy and those with a past history of bleeding gastric or duodenal ulcers were excluded. Participation in the study was voluntary and the invitation specified "a trial of new drug formulations." Of 189 patients invited to the study, 141 agreed to participate (74.6%). Nine could not do so because of inconvenience due to work, long distances to travel, or intercurrent illnesses; seven had moved from previous addresses and could not be contacted; two refused to participate; one patient had died; and one was due to be admitted to hospital during the period of the study.

The following clinical and laboratory features were noted at the time of the patient's attendance at the clinic: age, sex, rheumatoid arthritis classification, radiological stage and functional capacity (Steinbrocker *et al.*, 1949), duration of disease, articular index of joint tenderness (Ritchie *et al.*, 1968), presence or absence of subcutaneous nodules, serum albumin and globulin, haemoglobin, erythrocyte sedimentation rate, and rheumatoid and antinuclear factor titres. The patients were then asked to assess their degree of joint pain (initial pain rating) in terms of five subjective grades—namely, nil, mild, moderate, severe, and very severe. A printed sheet was provided and each patient instructed to record the degree of pain present according to these five grades just before retiring to bed each evening for the next 14 days (daily pain rating).

The patients were classified as having mild, moderate, or severe pain. The patients in each group were then allocated randomly to one of three treatment schedules: (1) placebo (calcium lactate one tablet four times daily), (2) prednisone 5 mg three times daily, and (3) aspirin (enteric coated) 975 mg four times daily. All other antirheumatic drugs taken before the trial were stopped for the duration of the study. It was carefully explained to each patient that if they had to discontinue the tablets the reason for doing so was to be recorded on the questionnaire provided (see Appendix). At the end of the trial the patients were asked to rate the drug according to five global scores of satisfaction—namely, totally ineffective, ineffective, moderately effective, effective, or highly effective—and then return the completed data charts by post in the stamped, addressed envelope provided.

For statistical purposes the five grades of pain were allocated the numerical scores of one to five where nil = 1, mild = 2,

Centre for Rheumatic Diseases, and University Department of Medicine, Glasgow Royal Infirmary, Glasgow

P. LEE, M.B., M.R.A.C.P., Senior Registrar in Rheumatology
J. WEBB, M.B., M.R.A.C.P., Lecturer in Rheumatology
W. W. BUCHANAN, M.D., F.R.C.P., Professor of Medicine

Department of Biomathematics, University of Oxford, Oxford
J. A. ANDERSON, M.A., D.PHIL. (OXON.), Lecturer in Medical Statistics

moderate = 3, severe = 4, and very severe = 5. Similarly the five global grades of satisfaction were scored from one to five where totally ineffective = 1, ineffective = 2, moderately effective = 3, effective = 4, and highly effective = 5.

STATISTICAL METHOD

The three variables that summarize the patient's response to treatment are the patient's assessment of his satisfaction with his treatment, the number of days (if any) withdrawn from the trial, and the average treated pain rating (A.T.P.R.). This last observation is the average of the recorded daily pain ratings: missing values were simply ignored. The procedure tends to reduce the levels of A.T.P.R. since the missing values correspond to days withdrawn from the trial, presumably with high pain scores.

The simplest comparisons between the three treatments on the basis of these measurements would be given by analyses of variance. It is known, however, that the patients start their treatments in very different states of health, so to make the analysis more precise adjustments were made to allow for these differences between the patients. A priori, it was thought that the initial pain rating (I.P.R.) would provide a very useful measure of the patient's health state at the beginning of the trial, but it was thought that some of the other observations (noted above) made as the patient entered the trial might be informative. In fact, once the I.P.R. had been allowed for, the contribution of the other observations was negligible. The differences between the three treatments were thus assessed by analyses of covariance, using the initial observations on the patients as concomitant variables (Cochran and Cox, 1957).

Results

Of the 141 patients who participated in the trial, 128 (90.8%) returned their completed pain charts. Of the 13 patients who failed to return the charts, one was unable to continue the trial because of the death of her husband and another developed an acute psychosis while on prednisone. The remaining 11 patients failed to complete the trial for unknown reasons. Five of these patients were on placebo, five on aspirin, and three were taking prednisone.

The mean I.P.R.s in the three groups (placebo 2.95, prednisone 3.00, and aspirin 2.88) are very similar, indicating a good randomization (table I). The differences in mean A.T.P.R. between aspirin and placebo, prednisone and placebo, and

TABLE I—Mean Initial Pain Rating (I.P.R.), Mean Average Treated Pain Rating (A.T.P.R.), and Mean A.T.P.R. Adjusted for Initial Pain Score for Patients in Each of Three Treatment Groups. S.E. (0.18) is S.E. of Difference between any Two Treatment Groups

Drug	No. of Patients	Mean I.P.R. (S.D.)	Mean A.T.P.R. (S.D.)	Mean A.T.P.R. Adjusted for I.P.R. (S.E.)
Placebo ..	41	2.95	3.47	3.47
Aspirin ..	42	2.88	3.01	3.04
Prednisone ..	45	3.00	2.56	2.53
		(1.0)	(0.83)	(0.18)

prednisone and aspirin were 0.46, 0.91, and 0.45. The differences between the three means were statistically significant at the 0.1% level in the straightforward F test in the analysis of variance. The inclusion of the I.P.R. as a covariate gave an analysis of covariance whose residual mean square was significantly smaller than that of the original analysis of variance ($P < 0.001$). Again, the between treatment differences in mean A.T.P.R. adjusted for I.P.R. were highly significant ($P < 0.001$), but the actual values of the adjusted mean A.T.P.R. were similar to the unadjusted means. This was because the mean I.P.R.s in the three groups were so similar. Except for

the articular index the effects of the other pretreatment observations (concomitant variables), having allowed for the initial pain score, were not statistically significant. The effect of the articular index was just statistically significant ($P < 0.05$) but the actual improvement in the precision of the analysis was slight (the residual mean square fell from 0.70 to 0.67).

Patients withdrew from the trial either for reasons related to pain or to side effects or both (table III). Because of difficulty in differentiating between pain and discomfort caused by side effects in many cases, all patients withdrawn from the trial

TABLE II—Numbers of Patients Withdrawn from Trial in Relation to Initial Pain Groups for Each Drug

Drug	Pain Severity			Total	χ^2
	Severe	Moderate	Mild		
Placebo					
Withdrawn ..	6	13	9	28	0.3 (N.S.)
Total ..	10	17	14	41	
Aspirin					
Withdrawn ..	4	8	2	14	2.8 (N.S.)
Total ..	10	17	15	42	
Prednisone					
Withdrawn ..	1	3	2	6	0.2 (N.S.)
Total ..	12	18	15	45	

N.S. = Not significant.

were considered under one class. Of the 128 patients included in the trial, 48 (37.5%) had withdrawn before completion of the trial period, and when these patients are separated into their various treatment groups it can be seen that there are considerable differences between the groups. Twenty-eight patients out of 41 (68.3%) on placebo, 14 out of 42 (33.3%) on aspirin, and six out of 45 (13.3%) on prednisone withdrew from the study. The chi-square test criterion for the differences between these proportions is 25.7 with 2 D.F., which is significant at the 0.1% level.

The question then arises whether the initial pain group of a patient has any bearing on whether or not he is likely to withdraw from the trial. The severe and very severe initial pain groups were amalgamated (table II) leaving three categories of initial pain. The differences between the proportions withdrawing in each of these groups were tested for each treatment separately by the chi-square criterion. None of the differences was statistically significant at the 5% level, leading to the rather surprising conclusion that the initial pain state of a patient does not appear to affect his chance of withdrawing from the test. Another way of assessing differences between the treatments in respect of days withdrawn from the trial is to compare the mean number of days withdrawn on each drug. These (table III) are statistically significant ($P < 0.0001$) on the analysis of variance F

TABLE III—Numbers of Patients Withdrawn from Trial and Mean Number of Days Withdrawn in each Treatment Group

Initial Pain Rating	Placebo		Aspirin		Prednisone		Total	
	With-drawn	Total No. in Group	With-drawn	Total No. in Group	With-drawn	Total No. in Group	With-drawn	Total No. in Group
Very severe	3	5	1	3	1	3	5	11
Severe ..	3	5	3	7	0	9	6	21
Moderate ..	13	17	8	17	3	18	24	52
Mild ..	9	14	2	15	2	15	13	44
Total ..	28 (68.3%)	41	14 (33.3%)	42	6 (13.3%)	45	48 (37.5%)	128
Mean No. of days with-drawn ..	6.4 (S.D. 4.2, S.E. 0.7)		3.0 (S.D. 4.2, S.E. 0.7)		1.3 (S.D. 4.2, S.E. 0.7)			

test. When the I.P.R. is a covariate in the corresponding analysis of covariance its effect is not statistically significant,

which reinforces the finding that the I.P.R. is not related to withdrawal from the trial.

The mean patient satisfaction ratings for each of the three drugs, placebo, aspirin, and prednisone were 1.70, 2.36, and 3.68 respectively (table IV). The differences between

TABLE IV—Mean Satisfaction Scores for Placebo, Aspirin, and Prednisone

	Mean Satisfaction Scores	
Placebo ..	1.70	} S.D. 1.26, S.E. 0.17 (approx. same for all groups)
Aspirin ..	2.36	
Prednisone ..	3.68	

the means of the three treatment groups (placebo and aspirin 0.66, placebo and prednisone 1.98, and aspirin and prednisone 1.32) were highly statistically significant ($P < 0.001$) on the analysis of variance F test. The satisfaction rating was analysed using the I.P.R. as covariate, but, surprisingly, its effect was not significant. The satisfaction rating was, however, significantly related to the average treated pain score ($P < 0.001$).

Discussion

In a review of clinical trials of indomethacin, O'Brien (1968) showed that improvement was more evident in those trials which did not include objective measurements. This could reflect the lower sensitivity of objective criteria in the assessment of patients with rheumatoid arthritis (Hart and Huskisson, 1972) and does not per se invalidate the use of subjective criteria. Indeed in our experience (Deodhar *et al.*, 1973) the most sensitive criteria in clinical trials of antirheumatic drugs have been the patient's assessment of joint pain and the articular index of joint tenderness, and the least sensitive indices have been objective measurements such as isotope "uptake" in the joints.

As Beecher (1959) so aptly stated: "pain is measured in terms of its relief," and it is pain which is the main complaint of the patient suffering from rheumatoid arthritis. Since anti-inflammatory analgesics do not prevent joint destruction and crippling deformities, the main concern of the doctor is relief of pain, and the clinical therapeutic benefit of an antirheumatic drug is essentially based on its analgesic properties. To our minds a measure of this effect should be an essential part of all trials of antirheumatic drugs.

Numerous studies in trials of analgesic drugs testify to the validity of using the patient's pain response or relief of pain in assessing their efficacy compared to placebo and to established analgesic preparations (Lee, 1942; Keele, 1948; Hewer *et al.*, 1949; Houde and Wallenstein, 1953; Lasagna and Beecher, 1954; Houde and Wallenstein, 1956; Houde *et al.*, 1960). Several workers have found that intelligent co-operative patients are able to give consistent reports on the course of their pain (Keele, 1948; Hewer *et al.*, 1949), and in a trial of pain relief using morphine in different doses, Keats *et al.* (1950) found the degree of error in the patient's assessments to be about 10%. Other workers (Lee, 1942; Houde and Wallenstein, 1953) encountered great difficulties in having the patient record their own pain charts—many of the charts being incomplete, wrongly filled in, or even lost. We have been impressed by the co-operation of our patients; only 13 of the 141 (9.2%) who agreed to participate failed to return their charts. The subjects were selected merely on their availability and willingness to co-operate. Most patients with rheumatoid arthritis have suffered continuous pain for years, and are only too willing to co-operate in a study which ultimately will be to their benefit. Of the 189 patients who were invited to participate in the study only two (1.1%) refused to take part.

The charts were made as simple as possible and the procedure

was carefully explained to each patient individually. We considered whether the patient should record his assessment of daily pain on separate sheets as advocated by Jacobsen (1965), or on the same sheet where he would be aware of the previous day's recordings. We chose the latter method because we thought it might help the patient to compare his pain with the previous day, and because separate sheets would lead to a greater risk of lost charts.

No ideal method has yet been devised for recording pain. Copeman (1950) used a percentage system whereby all patients began at 100% and moved up or down the scale according to progress. The visual analogue scale has recently been revived in psychiatric assessment but this method, like the percentage scale, has no descriptive limits and different patients with the same severity of pain may express their pain on widely varying divisions of the scales (Hart and Huskisson, 1972). This may be overcome to some extent in the visual analogue scale by "calibrating" the scale during the study (Berry and Huskisson, 1972). A scale of pain severity which we have used in the present study overcomes the problem of descriptive limits to some extent, but difficulties arise in the calculation of the response in that it has been assumed that the divisions in the scale are of equal length. Several workers (Houde and Wallenstein, 1953; Lasagna, 1960) have pointed out that the steps in such a graded scale may not in fact be of equal value and Lasagna (1960) was able to show that a fall in pain from "severe" to "moderate" was on the average the most important to the patient and a change from "slight" to "none" the least important. Nevertheless, Houde and Wallenstein (1953) concluded that even if the pain categories did not bear a linear relation to one another they do in fact represent appreciable differences in sensation, and showed that such data could be used, showing significant differences with the chi-square test, without assuming linearity.

Our numerical scale used five units to represent the five descriptive grades of pain: nil, mild, moderate, severe, and very severe. Possibly the results would be more accurate if each category of pain severity or its relief were weighted when assigning a scale and a set of numbers such as 0, 1.5, 3.0, 6.0, and 9.0 might be more appropriate. Studies done in educational testing suggest that it makes little difference (Wilks, 1938). Hart and Huskisson (1972) thought scales using only three or four grades of pain to be too insensitive to differentiate minor effects of drugs and have used a nine-point scale which produced better discrimination (Huskisson *et al.*, 1970). Nevertheless their four-point scale did show changes with aspirin, ibuprofen, and placebo, and Houde and Wallenstein (1953) when using a scale of one to five, found sufficiently sharp differentiations between aspirin, morphine, and placebo. We have adopted a five-point analogue scale for the present study mainly for the sake of simplicity, and also since too rigid an interpretation of the subjective analgesic scales can only contribute to the confusion in evaluating drug effects. Whether one uses an adjectival or numerical scale with the same number of divisions to delineate a subjective response makes no difference to the outcome (Beecher, 1959). Stevens and Galanter (1957) have shown in graphing category scales for loudness that the curve for the adjective scale falls eventually only a little below the numerical scale.

The assessment of analgesics when using the subjective methods has largely been applied to the relief of pain of post-operative trauma and of cancer origin. Since the appraisal of any drug must necessarily be made in the patient with the disease for which it was intended (Beecher, 1959), the results of such studies may not necessarily apply to patients such as those in the present study, all of whom suffered from rheumatoid arthritis. All had had a long experience of pain and treatment with a variety of antirheumatic drugs, and this might have been expected to have had some influence on the results over and above the effect of the initial pain rating; but this was not substantiated in the analysis of covariance. Once the I.P.R. has been allowed for, all the other initial readings have a negligible effect. Adaptation to pain does occur (Beecher, 1957)

but no systematic studies on this subject to our knowledge have been carried out. A decreased pain sensitivity with increased pain perception and reaction thresholds have been found with increasing age (Chapman and Jones, 1944; Sherman and Robillard, 1960), and older people tend to respond to drugs and placebos with a greater frequency of beneficial response than do younger age groups (Lasagna *et al.*, 1954; Bender, 1964; Bellville *et al.*, 1971); but there was no suggestion of these phenomena in the present study.

Of the initial observations made on each patient, only the articular index score was significantly correlated with the average treated pain rating having allowed for the initial pain rating. The inclusion of the initial pain score in the analysis significantly increased its precision, and both change in pain intensity and estimates of relief are dependent on initial pain intensity (Gruber and Batisti, 1963). Nevertheless, the actual improvement in precision was only slight with the additional inclusion of the articular index score. It is of great interest that the global satisfaction rating for each drug is significantly related to the average treated pain score, and highly significant differences were noted for prednisone, aspirin, and placebo in both cases. Dick *et al.* (1970b) found patient preference to be as good as several other criteria in assessing the effectiveness of drugs, although Hart and Huskisson (1972) have pointed out that preference is compounded of many factors and a patient may prefer a drug because of its freedom from side effects rather than because of its analgesic potency. Taylor *et al.* (1971) found patients' drug preference to be less sensitive than a scale of pain relief. Surprisingly, however, none of the initial observations, including the initial pain rating, are significantly related to the final satisfaction rating. A reasonable explanation of this is that when a patient assesses his overall satisfaction he automatically discounts his initial state. Clearly from our result the more effective the drug the fewer patients were likely to withdraw from the trial. From the analysis of the number of patients withdrawn from each initial pain group for each drug, it is concluded that the initial pain state of a patient does not appear to affect materially his chance of withdrawing.

One patient developed an acute psychotic reaction during the trial while taking prednisone. Mental instability was not a feature in the patient's past history and she made a satisfactory recovery after electroconvulsive therapy. Though other factors are likely to be involved in this case, clearly one must be cautious when proceeding with corticosteroid therapy, even in moderate dosage as used in the current trial.

The inclusion of placebo in the present study raises ethical problems. A placebo is by no means therapeutically ineffective (Beecher, 1959), and the impressive initial response to many new medications is undoubtedly due to this effect. In our opinion it is essential to include placebo to avoid the conclusion: drug A = drug B = 0. In this study it was explained to the patient that the drug may not be effective, and provision was made in the experimental design to terminate the trial if the

pain became unbearable. Since none of the currently available anti-inflammatory drugs has been shown to fundamentally alter the course of rheumatoid arthritis, we consider it justifiable to use a placebo.

We wish to express gratitude for financial support from the Arthritis and Rheumatism Council for Research in Great Britain.

Requests for reprints should be addressed to: Dr. P. Lee, The Centre for Rheumatic Diseases, 35 Baird Street, Glasgow G4 0EH.

References

- Beecher, H. K. (1957). *Pharmacological Reviews*, 9, 59.
 Beecher, H. K. (1959). *Measurement of Subjective Response*, Chapter 2. New York, Oxford University Press.
 Bellville, J. W., Forrest, W. H., Miller, E., and Brown, B. W. (1971). *Journal of the American Medical Association*, 217, 1835.
 Bender, A. D. (1964). *Journal of the American Geriatrics Society*, 12, 114.
 Berry, H., and Huskisson, E. C. (1972). Cited by Hart, F. D., and Huskisson, E. C. (1972).
 Boardman, P. L., and Hart, F. D. (1967). *British Medical Journal*, 4, 264.
 Chapman, W. F., and Jones, C. M. (1944). *Journal of Clinical Investigation*, 23, 81.
 Cochran, W. G., and Cox, G. M. (1957). *Experimental Designs*, p. 82. London, Wiley.
 Collins, A. J., and Cosh, J. A. (1970). *Annals of the Rheumatic Diseases*, 29, 386.
 Collins, K. E., *et al.* (1971). *Annals of the Rheumatic Diseases*, 30, 401.
 Copeman, W. S. C. (1950). *British Medical Journal*, 2, 849.
 Cosh, J. A., and Ring, E. F. J. (1967). *Journal de Radiologie, d'Electrologie et Archives d'Electricité Médicale*, 48, 84.
 Deodhar, S. D., Dick, W. C., Hodgkinson, R., and Buchanan, W. W. (1973). *Quarterly Journal of Medicine*. In press.
 Dick, W. C., *et al.* (1970a). *Annals of the Rheumatic Diseases*, 29, 135.
 Dick, W. C., Grayson, M. F., Woodburn, A., Nuki, G., and Buchanan, W. W. (1970b). *Annals of the Rheumatic Diseases*, 29, 643.
 Gruber, C. M., and Batisti, A. (1963). *Clinical Pharmacology and Therapeutics*, 4, 172.
 Hart, F. D., and Clark, C. J. M. (1951). *Lancet*, 1, 775.
 Hart, F. D., and Huskisson, E. C. (1972). *Lancet*, 1, 28.
 Hewer, A. J. H., Keele, C. A., Keele, K. D., and Nathan, P. W. (1949). *Lancet*, 1, 431.
 Hollander, J. L., Stoner, E. K., Brown, E. M., jun., and deMoor, P. (1951). *Journal of Clinical Investigation*, 30, 701.
 Houde, R. W., and Wallenstein, S. L. (1953). *Drug Addiction and Narcotics Bulletin, Appendix C*, p. 417.
 Houde, R. W., and Wallenstein, S. L. (1956). *Drug Addiction and Narcotics Bulletin, Appendix B*, p. 1383.
 Houde, R. W., Wallenstein, S. L., and Rogers, A. (1960). *Clinical Pharmacology and Therapeutics*, 1, 163.
 Huskisson, E. C., Shenfield, G. M., Taylor, R. T., and Hart, F. D. (1970). *Rheumatology and Physical Medicine*, Suppl. p. 88.
 Jacobsen, M. (1965). *British Journal of Psychiatry*, 111, 545.
 Keats, A. S., Beecher, H. K., and Mosteller, F. C. (1950). *Journal of Applied Physiology*, 1, 34.
 Keele, K. D. (1948). *Lancet*, 2, 6.
 Lansbury, J. (1958). *Arthritis and Rheumatism*, 1, 505.
 Lansbury, J. (1966). In *Arthritis and Allied Conditions, Textbook of Rheumatology*, ed. J. L. Hollander, 7th ed. Philadelphia, Lea and Febiger.
 Lasagna, L. (1960). *Annals of the New York Academy of Sciences*, 86, 28.
 Lasagna, L., and Beecher, H. K. (1954). *Journal of the American Medical Association*, 156, 230.
 Lasagna, L., Mosteller, F., and von Felsinger, J. M. (1954). *American Journal of Medicine*, 16, 770.
 Lee, L. E. (1942). *Journal of Pharmacology*, 75, 161.
 Lee, P., Baxter, A., Dick, W. C., and Webb, J. (1973). Submitted for publication.
 Mainland, D. (1967). *Arthritis and Rheumatism*, 10, 71.
 Medical Research Council and Nuffield Foundation Joint Committee (1960). *Annals of the Rheumatic Diseases*, 19, 331.
 O'Brien, W. M. (1968). *Clinical Pharmacology and Therapeutics*, 9, 94.
 Ritchie, D. M., *et al.* (1968). *Quarterly Journal of Medicine*, 37, 393.
 Ropes, M. W., Bennett, G. A., Cobb, S., Jacox, R., and Jessar, R. A. (1959). *Annals of the Rheumatic Diseases*, 18, 49.
 Sherman, E. D., and Robillard, E. (1960). *Canadian Medical Association Journal*, 83, 944.
 Steinbrocker, O., Traeger, C. H., and Batterman, R. C. (1949). *Journal of the American Medical Association*, 140, 659.
 Stevens, S. S., and Galanter, E. H. (1957). *Journal of Experimental Psychology*, 54, 377.
 Taylor, R. T., Huskisson, E. C., and Hart, F. D. (1971). *Abstracts of the VII European Rheumatology Congress*, Section 18, Abstract 9.
 Webb, J., Downie, W. W., Dick, W. C., and Lee, P. (1973). *Scandinavian Journal of Rheumatology*. In press.
 Weiss, T. E., Maxfield, W. S., Murison, P. J., and Hildago, J. V. (1966). *Southern Medical Journal*, 59, 484.
 Whaley, K., *et al.* (1968). *Clinical Science*, 35, 547.
 Wilks, S. S. (1938). *Psychometrika*, 3, 23.
 Wright, V. (1959). *Clinical Science*, 18, 17.

APPENDIX—Questionnaire Provided for Completion if Patients Withdrew From Trial

IF YOU DO NOT COMPLETE 14 DAYS' TREATMENT, PLEASE EXPLAIN WHY.

I stopped the tablets because:

- | | |
|---|--------|
| (1) They did not help me and I was having continuously severe pain. | YES/NO |
| (2) The tablets upset me | YES/NO |
| (3) Did you experience any side effects from the tablets? | YES/NO |
| (4) If "yes," what did you feel wrong? | |

WHEN COMPLETE PLEASE RETURN BY POST IN STAMPED ADDRESSED ENVELOPE PROVIDED.