

activity levels which was common to all the subjects rather than to the absolute levels of physical condition. Those who had been in the best condition before operation were not fully restored by two weeks rehabilitation, and those who had been in poor condition before operation improved over and above their initial condition. The results of Saltin *et al.* (1968) were similar except that the improvement due to rehabilitation was very much greater (a decrease of about 40 beats/min in standardized heart rate) owing to the severity and length of the programme (eight weeks strenuous interval training at near maximum work rates). As in the present study the improvement seemed to be related to the degree of change in activity levels and not to the absolute levels of physical condition.

Conclusion.—In principle, prophylactic training for elective surgery is attractive. Any measure which may reduce the impact of surgery and accelerate recovery is of importance. This study suggests that good physical condition at the time of hospital admission does not prevent deterioration in physical condition during bed rest or influence the benefits to be derived from rehabilitation. These conclusions apply to only one aspect of cardiorespiratory function.

This investigation formed part of a project supported by a grant from the Nuffield Foundation. We thank the staff of Harlow Wood Orthopaedic Hospital, near Mansfield, for their encouragement, the provision of facilities, and their co-operation. We are particularly grateful to Mr. N. J. Barton, Mr. J. P. Campbell, Mr. J. S. Hopkins, Mr. J. P. Jackson, Dr. K. Lloyd-Jones, and Mr. W. Waugh for permitting us to study patients admitted under their care.

References

- Adolfsson, G. (1969). *Scandinavian Journal of Rehabilitation Medicine* 1, 14.
 Alvik, I. (1966). In *Physical Activity in Health and Disease*, ed. K. Evang and K. Lange Anderson. Baltimore, Williams and Wilkins.
 Astrand, P. O. (1956). *Physiological Reviews*, 36, 307.
 Bassey, E. J., *et al.* (1973). *Cardiovascular Research*, 7, 588.
 Bassey, E. Joan and Goldsmith, R. (1974). *Scandinavian Journal of Rehabilitation Medicine*. In press.
 Cotes, J. E. (1971). *Scandinavian Journal of Respiratory Diseases*, Suppl. No. 77, p. 123.
 Miller, P. B., Johnson, R. L., and Lamb, L. E. (1965). *Aerospace Medicine*, 36, 1077.
 Rodahl, K., *et al.* (1967). In *Nutrition and Physical Activity*, ed. G. Blix, p. 107. Stockholm, Almqvist and Wiksell.
 Saltin, B., *et al.* (1968). *Circulation*, 38, Suppl. No. 7, p. 1.
 Taylor, H. L., *et al.* (1949). *Journal of Applied Physiology*, 2, 233.

Simple Analgesics for Arthritis

E. C. HUSKISSON

British Medical Journal, 1974, 4, 196-200

Summary

A series of experiments has been carried out with single doses of simple analgesics in patients with rheumatoid arthritis using a consistent polyad design. This method proved to be both valid and useful. Pain relief scores were a better measure of the effectiveness of analgesics than preference.

Aspirin, Codis, and Distalgic were the most effective analgesics tested, with paracetamol, pentazocine, and Ciba 44,328 intermediate between these agents and placebo. Placebo given after an active analgesic was more effective than when given before; this phenomenon was not abolished by telling the patients that apparently identical tablets were, in fact, different or by making them different in colour. The effectiveness of soluble placebo depended on its colour, red being the most effective.

Introduction

It is conventional to supplement regular anti-inflammatory drug therapy for rheumatoid arthritis with single doses of simple analgesics, the patient being instructed to take some tablets when pain is uncontrolled. If analgesics are not provided many patients will obtain them without prescription. In a previous study (Huskiison and Hart, 1972 a) patients with rheumatoid arthritis were found to take an average of 9.2 tablets daily, of which 4.2 were simple analgesics taken "on demand"; the number of tablets taken was related to the patient's pain threshold as well as to the severity of the disease.

The object of this study was to find a suitable experimental design for testing simple analgesics in patients with arthritis and to use this design to test some of the available drugs. Some conditions were imposed on the experiment in advance. Firstly, since there are a large number of analgesics from which to choose the method had to be capable of testing large numbers of compounds. Secondly, since simple analgesics are given in single doses for uncontrolled pain they had to be tested in this way. Thirdly, the design had not to require an impossibly large sample size, and it was anticipated that about 24 patients would be readily available for such an experiment. This combined with the need to test large numbers of compounds made a cross-over design essential. Fourthly, the method had to be directed towards the measurements of effectiveness in terms of pain relief; toxicity could be studied in separate, long-term experiments designed for the purpose which would give the opportunity for the demonstration of other important phenomena such as interactions with regular treatment.

Three experiments are reported here, the first comparing analgesic tablets, the second soluble analgesics in different colours, and the third a new analgesic with a standard and placebo. The trials were designed to take advantage of the consistent polyad method described by Kendall (1963) and used previously (Huskiison and Hart, 1972 b) in a trial of single doses of anti-inflammatory drugs. This method is ideal for multiple comparisons because every possible pair of treatments is directly compared and the variability between first and last is thereby minimized.

Methods

Thirty patients were admitted to the first trial, 24 to the second, and 24 to the third. All had definite or classical rheumatoid arthritis (A.R.A. criteria; Ropes *et al.*, 1959), with pain of sufficient severity to require "on-demand" analgesics at least once a day.

First Trial.—Four treatments were compared in the first trial, Distalgesic (paracetamol 325 mg and dextropropoxyphene 32.5 mg per tablet) 2 tablets; paracetamol (Panadol, 500 mg per tablet) 2 tablets; pentazocine (Fortral, 25 mg per tablet) 2 tablets; and placebo 2 tablets. The double-placebo method was used to ensure that the trial was double blind and each patient was therefore given four tablets in each dose. The patients were asked to take all four tablets in place of their usual on-demand analgesics on the first day of the study; four more tablets were supplied for day 2 but the third day was a rest day. Test tablets were supplied for days 4, 5, 7, 8, 10, 11, 13, 14, 16, and 17 of the study. Each patient compared all six possible treatment pairs. Treatment order was randomized and balanced so far as possible. Returned bottles were checked to ensure that all the tablets had been taken. Regular anti-inflammatory therapy continued unchanged. On-demand analgesics were permitted except during the six hours after administration of the test analgesics. To avoid additive effects none of the trial drugs were allowed during the experiment. For six hours after administration of the test analgesics the patients recorded their pain relief score. Pain relief was graded as none, slight, moderate, or complete. On the second day of each comparison they noted their preference for one or other treatment of the pair.

Second Trial.—Four soluble analgesics were compared in the second trial, Codis (aspirin 500 mg and codeine 8 mg per tablet) 2 tablets; aspirin (Solprin, 300 mg per tablet) 2 tablets; Distalgesic 2 tablets; and placebo 2 tablets. The tablets were identical in shape and size but each was made in four colours, red, blue, green, and yellow. Four colour codes were used to ensure that each treatment in a pair comparison appeared different. For a particular patient the same colour code was used throughout. Thus a patient given red Distalgesic in the first comparison received red Distalgesic throughout; the next patient received green Distalgesic throughout. The order of treatment and the colours used were randomized so that about the same number of patients received each treatment in each order and each colour. The conduct of the experiment was identical with that of the first trial except that the patients dissolved the tablets in water.

Third Trial.—Three treatments were compared in this trial: a new analgesic, Ciba 44,328 (the dextrorotary isomer of 2-(dimethylaminomethyl)-1,2-diphenyl-3-butene-2-ol propionate in the maleate form, a compound related to dextropropoxyphene, 50 mg per tablet) 2 tablets; paracetamol 2 tablets; and placebo 2 tablets. Non-soluble, identical white tablets were used and the design of the trial was similar to that of the first and second trials except that it lasted only nine days. Because there were only three treatments there were only three comparisons instead of six, and each drug was given only twice instead of three times.

STATISTICAL

The results were analysed in two ways, on the basis of preferences and by pain relief scores. The pain relief scores for three administrations of each drug were added to produce

a total score at each hour after administration. This raises the problem of multiple comparisons. Scores on which a comparison of two drugs is based are taken from five-pair comparisons, and this must increase the probability of type-1 error. Miller (1966) provided a formula from which it is possible to calculate this error; in the five comparisons using a significance level of $P < 0.05$ the probability of finding a difference where none exists is $1 - (0.95)^5$ or 23%. When the significance level is reduced to $P < 0.01$ the probability of type-1 error becomes an acceptable 6%. Though it is uncertain that as many as five comparisons are involved a level of $P < 0.01$ was chosen as a suitable level of significance for experiments using multiple comparisons.

Differences within patients in total pain relief scores at each hour were analysed using the Friedman two-way analysis of variance by ranks; Wilcoxon's signed ranks matched pairs test was then used to identify significant differences between treatments. The Kruskal-Wallis one-way analysis of variance was used to test differences between groups of patients in response to coloured placebo. These methods were described by Siegel (1956).

Preferences were analysed using the sign test (Siegel, 1956) as well as by the consistent triad method described by Kendall (1963).

Results

FIRST TRIAL

Twenty-three patients completed the first trial. The mean total relief scores are shown in fig. 1 and the results of statistical tests applied to them are given in table I. Even using the 1% significance level there were clear differences in favour of all active drugs over placebo and in favour of Distalgesic over pentazocine and paracetamol. There was no significant difference between pentazocine and paracetamol.

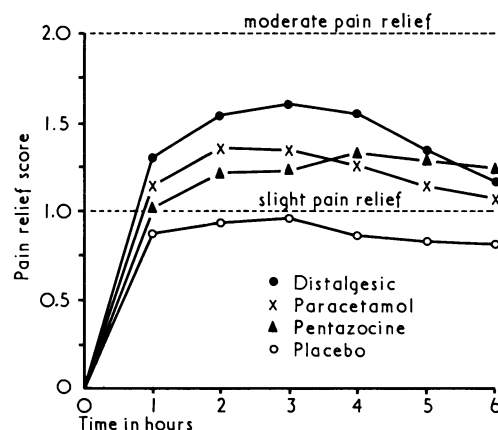


FIG. 1—First Trial. Mean pain relief scores for single doses of Distalgesic, pentazocine, paracetamol, and placebo. Mean scores are derived from three administrations of each treatment to 23 patients completing trial.

TABLE I—Significance Levels for Statistical Tests Carried Out on Total Pain Relief Scores in First Trial

No. of Hours after Administration	Friedman 2-way Analysis of Variance	Wilcoxon Tests					
		Paracetamol v. Placebo	Pentazocine v. Placebo	Distalgesic v. Placebo	Paracetamol v. Pentazocine	Distalgesic v. Pentazocine	Distalgesic v. Paracetamol
1	$P < 0.05$	$P < 0.02$	N.S.	$P < 0.01$	N.S.	$P < 0.05$	N.S.
2	$P < 0.005$	$P < 0.01$	$P < 0.01$	$P < 0.01$	N.S.	$P < 0.05$	N.S.
3	$P < 0.005$	$P < 0.01$	$P < 0.05$	$P < 0.01$	N.S.	$P < 0.01$	$P < 0.05$
4	$P < 0.001$	$P < 0.01$	$P < 0.01$	$P < 0.01$	N.S.	$P < 0.02$	$P < 0.01$
5	$P < 0.005$	$P < 0.01$	$P < 0.01$	$P < 0.01$	N.S.	N.S.	$P < 0.05$
6	N.S.	$P < 0.01$	$P < 0.01$	$P < 0.01$	N.S.	N.S.	N.S.

N.S. = Not Significant ($P < 0.05$)

These results suggest that the method is sufficiently sensitive to show clinically significant differences, though it should be noted that these drugs produced on average only slight to moderate pain relief, and the effectiveness of placebo was greater than the difference between active drugs and placebo. Mean pain relief scores taken from only the direct comparisons of each pair of compounds were qualitatively similar to those shown in fig. 1 but the differences were smaller; this suggests that it is both valid and useful to combine the scores for three administrations of each drug.

In this trial preference was based on differences in total pain relief scores when it was not directly stated. Analysis of preferences using Kendall's method showed a statistically highly significant coefficient of agreement between patients ($\chi^2 = 20.9$; $P < 0.005$). Using only the results from 10 patients whose preferences were consistent there were statistically significant differences between all three active drugs and placebo (sign test, $P < 0.05$) but not between any of the active drugs. All preferences whether consistent or inconsistent are shown in table II. It should be noted that the results are inconsistent; Distalgesic was not shown to be superior to placebo though it was superior to pentazocine, which was itself superior to placebo. Individual patient's preferences were consistent in only 10 of the 23 cases (43%). One reason for the failure of preference to measure pain relief was noted: some patients felt sick after taking pentazocine and preferred another analgesic or placebo even though it provided less pain relief.

Placebo was found to be more effective when given after active medication than when given before (fig. 2).

TABLE II—Preferences Expressed by 23 Patients Completing First Trial

Preference	No. of Patients					
Pentazocine	17	14	5	10	7	12
Paracetamol			13	15	12	
Distalgesic	2	4	7	15	12	
Placebo	4	5	3	3	1	4
None						
P (sign test) ..	0.004	0.03	>0.1	0.04	>0.1	>0.1

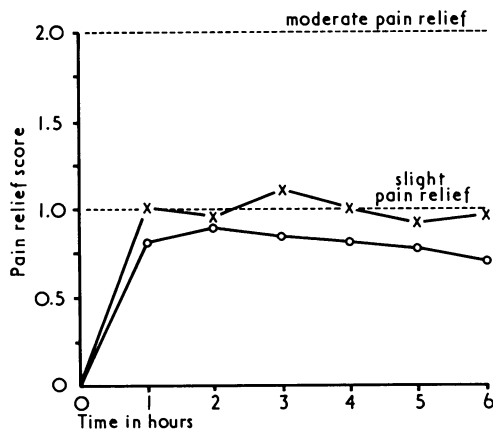


FIG. 2—First trial. Mean pain relief scores for single doses of placebo given either before (o) or after (x) active analgesics using identical formulation.

SECOND TRIAL

Twenty-two patients completed this trial. The mean total pain relief scores are shown in fig. 3 and the results of statistical tests applied to them are given in table III. There were clear differences favouring all three active drugs over placebo but no significant differences between the active drugs. Two treatments used in this trial (Distalgesic and placebo) were also used in the first trial; the effects were closely similar even though a different group of patients was used.

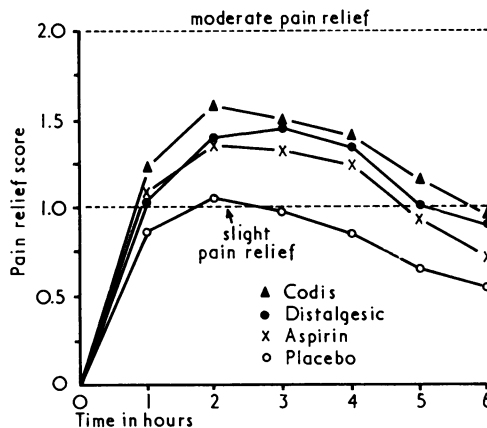


FIG. 3—Second trial. Mean pain relief scores for single doses of soluble Distalgesic, Codis, aspirin, and placebo. Mean scores are derived from three administrations of each treatment to 22 patients completing trial.

When the effects of four different colours of placebo were compared (fig. 4) it appeared that red placebo was as effective as any of the active drugs. The difference between the colours was statistically significant except at the first hour after administration (Kruskal-Wallis test: $P < 0.02$ at two and four

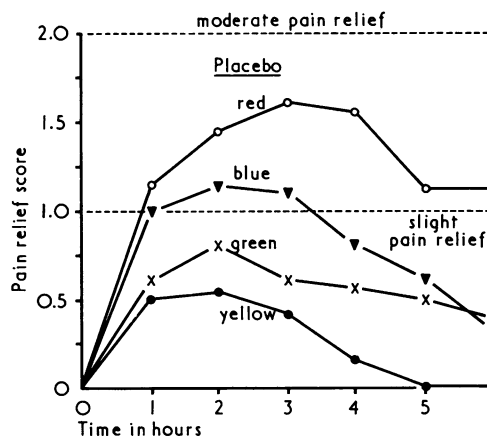


FIG. 4—Second trial. Mean pain relief scores for single doses of placebo in four different colours.

TABLE III—Significance Levels for Statistical Tests Carried Out on Total Pain Relief Scores in Second Trial

No. of Hours after Administration	Friedman 2-way Analysis of Variance	Wilcoxon Tests					
		Aspirin v. Placebo	Codis v. Placebo	Distalgesic v. Placebo	Aspirin v. Codis	Aspirin v. Distalgesic	Codis v. Distalgesic
1	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
2	$P < 0.05$	$P < 0.01$	$P < 0.01$	$P < 0.05$	N.S.	N.S.	N.S.
3	$P < 0.02$	$P < 0.01$	$P < 0.01$	$P < 0.01$	N.S.	N.S.	N.S.
4	$P < 0.01$	$P < 0.01$	$P < 0.01$	$P < 0.01$	N.S.	N.S.	N.S.
5	$P < 0.05$	N.S.	$P < 0.02$	$P < 0.02$	N.S.	N.S.	N.S.
6	N.S.	N.S.	N.S.	$P < 0.05$	N.S.	N.S.	N.S.

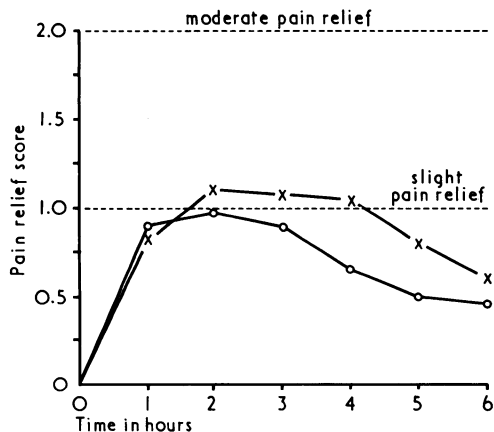


FIG. 5—Second trial. Mean pain relief scores for single doses of placebo given either before (o) or after (x) active analgesics using non-identical formulation.

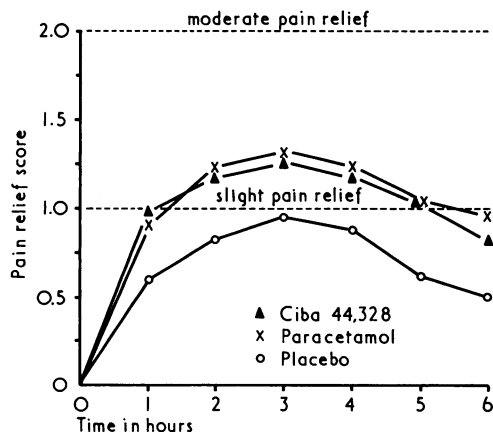


FIG. 6—Third trial. Mean pain relief scores for single doses of Ciba 44,328; paracetamol; and placebo. Mean scores are derived from two administrations of each treatment to 22 patients completing trial.

hours; $P < 0.05$ at three, five, and six hours). This difference was not seen with active drugs.

The difference in the effect of placebo given either before or after active analgesics persisted despite the use of different colours in each pair comparison (fig. 5).

In this trial preferences were analysed as expressed by the patients without reference to pain relief scores. With the use of Kendall's method a statistically significant coefficient of agreement was shown between patients ($\chi^2=14.5$; $P < 0.05$). Only eight patients, however, expressed consistent preferences, and using only the results from these none of the differences were statistically significant. All preferences whether consistent or inconsistent are shown in table IV.

TABLE IV—Preferences Expressed by 22 Patients Completing Second Trial

Preference	No. of Patients					
	1	2	3	4	5	6
Distalgesic	17	14	10	13	9	8
Codis			10	11	6	
Aspirin			10			
Placebo	4	6	2	1	3	5
None	1	2				
P (sign test) ..	0.008	>0.1	>0.1	>0.1	>0.1	>0.1

Only the difference between Distalgesic and placebo was statistically significant ($P < 0.01$), a curious paradox in view of the failure of this comparison to reach statistical significance in the first trial. Altogether 90% of all possible preferences were expressed; in the remaining 10% either the treatments were said to be equal or no preference was stated. Only 59% of the patients, however, expressed a preference in all six pair comparisons, so that there were many in which it was not possible to test consistency. Only eight of the 22 patients (36%) expressed consistent preferences.

THIRD TRIAL

Twenty-two patients completed this trial. The mean total pain relief scores are shown in fig. 6 and the results of statistical tests applied to them are given in table V. There were clear differences favouring both paracetamol and Ciba 44,328 when compared with placebo but no difference between the active drugs.

With the use of Kendall's method a statistically significant coefficient of agreement was shown between patients ($\chi^2=10.1$; $P < 0.05$). When only consistent preferences were analysed both paracetamol and Ciba 44,328 were superior to placebo (sign test, $P < 0.05$) but the difference between them was not statistically significant.

TABLE V—Significance Levels for Statistical Tests Carried Out on Total Pain Relief Scores in Third Trial

No. of Hours after Administration	Wilcoxon Tests		
	Ciba 44,328 v. Placebo	Paracetamol v. Placebo	Paracetamol v. Ciba 44,328
1	$P < 0.05$	N.S.	N.S.
2	$P < 0.02$	$P < 0.02$	N.S.
3	N.S.	$P < 0.02$	N.S.
4	$P < 0.05$	N.S.	N.S.
5	$P < 0.01$	$P < 0.05$	N.S.
6	$P < 0.01$	$P < 0.01$	N.S.

Discussion

This method of testing simple analgesics in rheumatoid arthritis proved to be satisfactory when pain relief scores were used to measure effectiveness. Preference was less satisfactory for three main reasons. Firstly, whereas pain relief scores produced consistent results for the same drug in different experiments using different subjects preferences did not. Secondly, preference did not necessarily reflect pain relief; in the second experiment 22% of preferences disagreed with the order of preference determined by pain relief scores. In some cases preference was not expressed even though pain relief scores were different; in others side effects from the more effective drug led to a preference for the less effective one. Thirdly, preferences were not expressed with sufficient regularity or consistency for this to be a reliable method of assessment. In the second experiment only 59% of the patients expressed all six preferences and only 36% expressed consistent preferences.

In the first experiment placebo was shown to be more effective when given after an active analgesic than when given before. This was observed by Sunshine *et al.* (1964) and also by Kantor *et al.* (1966), who showed a relationship between the effectiveness of a compound and that of a subsequent placebo. It may be surmised that a patient who associates pain relief with a tablet of a certain size, shape, and colour will obtain pain relief from an apparently identical tablet even though it is a placebo. Cromie (1963) recognized that this was a disadvantage and led to a reduction in the sensitivity of the method; since the effectiveness of placebo is increased the difference between placebo and active drugs is decreased. Cromie (1963) suggested that patients taking part in a double-blind experiment should be told that the treatments are different even though they appear identical; he also suggested the use of colour to ensure that the treatments are not identical. These experiments show that neither of these procedures abolishes the phenomenon. It is becoming increasingly difficult to make different treatments appear identical

without altering the bioavailability of the compounds (Garnham, 1974). It is also clear that the appearance of the medication must be considered in the design of experiments and in the interpretation of results. Effects of colour have been shown in other conditions, green being preferred for anxiety and yellow for depression (Schapira *et al.*, 1970).

The experimental design used in this study proved to be satisfactory for testing simple analgesics. There is reason to think that single-dose experiments may be more sensitive for the measurement of analgesic potency than studies based on multiple doses. Nuki *et al.* (1973) failed to show the effectiveness of pentazocine given regularly for one week to patients with rheumatoid arthritis, whereas in this study a single dose of pentazocine was clearly effective. If a compound produces its maximal effectiveness a few hours after ingestion its action is more likely to be detected by measurement made at that time rather than at the end of a week of multiple doses; the week must be compounded of many hours of little or no pain relief as well as the few hours of modest pain relief which these compounds provide.

The results of our three trials showed that in rheumatoid arthritis single doses of Distalgesic, Codis, and aspirin were about equally effective; pentazocine, paracetamol, and Ciba 44,328 were intermediate in effectiveness between these agents and placebo. Moertel *et al.* (1972), who studied patients with unresectable cancer, also found oral pentazocine (50 mg) and paracetamol (65 mg) similar and intermediate in effectiveness between aspirin (650 mg) and placebo. Kantor *et al.* (1966) in a study of patients with postoperative pain or fractures also found aspirin (600 mg) superior to pentazocine (50 mg).

The choice of a simple analgesic for patients with rheumatic conditions should be determined by the toxicity as well as

the effectiveness of available medicines. Lack of toxicity is a particularly important feature in view of the rather modest effects of these drugs; they produce only slight to moderate pain relief, whereas placebo is not much less effective. Distalgesic and aspirin appear to be the first choices for effectiveness; Distalgesic may have the advantage of causing fewer side effects but no formal comparison has been made. There is little justification for using codeine with aspirin, and oral pentazocine cannot be recommended because of its side effects. Ciba 44,328 does not merit addition to the already long list of available pain killers.

I am indebted to Dr. H. Wykeham Balme and Dr. F. Dudley Hart, whose patients I studied; Miss Jane Scott for help with the conduct of the trials and the statistics; Mrs. Hazel Lane and Mrs. Ann Darby-Dowman for statistical advice; and Dista Products Ltd., Winthrop Laboratories, Reckitt and Colman Ltd., and Ciba Laboratories for supplies of analgesics.

References

- Cromie, B. W. (1963). *Lancet*, 2, 994.
 Garnham, J. C. (1974). *Postgraduate Medical Journal*, 50, 114.
 Huskisson, E. C., and Hart, F. D. (1972 a). *British Medical Journal*, 4, 193.
 Huskisson, E. C., and Hart, F. D. (1972 b). *Practitioner*, 208, 248.
 Kantor, T. G., *et al.* (1966). *Clinical Pharmacology and Therapeutics*, 7, 447.
 Kendall, M. G. (1963). *Rank Correlation Methods*, 3rd edn. London, Griffin.
 Miller, R. G. (1966). *Simultaneous Statistical Inference*. New York, McGraw Hill.
 Moertel, C. G., *et al.* (1972). *New England Journal of Medicine*, 286, 813.
 Nuki, G., *et al.* (1973). *Annals of Rheumatic Diseases*, 32, 436.
 Ropes, M. W., *et al.* (1959). *Annals of Rheumatic Diseases*, 18, 49.
 Schapira, K., *et al.* (1970). *British Medical Journal*, 2, 446.
 Siegel, S. (1956). *Non-Parametric Statistics for the Behavioral Sciences*. New York, McGraw Hill.
 Sunshine, A., *et al.* (1964). *Clinical Pharmacology and Therapeutics*, 5, 699.

Inverse Relation between Serum IgG Concentration and Glucose and Xylose Absorption in Zambian African Adults

G. C. COOK

British Medical Journal, 1974, 4, 200-201

Summary

In three separate investigations in Zambian African adults in Lusaka an inverse association has been shown between serum IgG concentration and either glucose absorption rate from the jejunum (from a 200-mM solution) or xylose absorption (from a 25-g oral load). In developing areas of the world, where serum IgG concentrations are often high, this finding may have widespread nutritional implications.

Introduction

Subclinical malabsorption of monosaccharides, shown by abnormal xylose absorption and flat oral glucose tolerance

curves, is common in African adults in tropical Africa. Morphological abnormalities of the jejunal mucosa and the presence of intestinal parasites do not usually explain this (Cook *et al.*, 1969). Systemic bacterial infections, such as lobar pneumonia and pulmonary tuberculosis, have, however, been associated with a depression of glucose (Cook, 1971) and xylose (Cook, 1972) absorption in Zambian African patients. Furthermore, in two recent studies the jejunal absorption rate of glucose from a 200-mM solution (Cook, 1973) and the weight of xylose excreted after a 25-g oral load (Cook, 1974) have been shown to be inversely related to serum γ -globulin concentration in separate groups of well-nourished Zambian African adults.

The present investigation was undertaken to determine whether carbohydrate malabsorption is associated with a high concentration of one or more of the serum immunoglobulin fractions. Serum concentrations of immunoglobulins IgG, IgA, IgM, and IgD were determined in most of the subjects studied in two previous investigations (Cook, 1973, 1974), and correlations with glucose and xylose absorption were made. In an additional group of patients the relation between serum IgG concentration and xylose absorption was studied.

Department of Medicine, The University of Zambia, P.O. Box RW 110, Lusaka, Zambia
 G. C. COOK, M.D., F.R.C.P., Professor