

MEDICAL PRACTICE

Contemporary Themes

A defence of the small clinical trial: evaluation of three gastroenterological studies

J POWELL-TUCK, K D MACRAE, M J R HEALY, J E LENNARD-JONES, R A PARKINS¹

Emphasis has been placed on the importance of β , the type II error, in clinical trials and the corresponding necessity of including large numbers of subjects in a study.¹ Large trials are in fashion and small trials—even when they yield statistically significant results—are regarded with suspicion; a small trial which does not show a conventional statistically significant difference is often thought worthless. Interpretation of the small trial has increasingly been recognised to be dangerous, and the argument that lack of a significant difference between two treatments must mean that they have equal efficacy is recognised to be false. The need for large trials is undeniable in diseases such as cancer and cardiovascular problems where earlier trials have excluded large advantages for new treatments and small differences in outcome are still clinically and socially important. The unfortunate consequence of the emphasis on large trials in these settings is that small trials conducted in other diseases where only large differences in outcome are relevant have illogically become unacceptable in many people's minds.

With appropriate statistical analysis, including calculation of the confidence limits for the likely "true" difference, small controlled trials can provide clinically useful information. We present three such small clinical trials—none of which show conventional statistically significant results and none of which, given the present climate of opinion, would be likely to be accepted for publication in their own right; we discuss their analysis and show that important practical decisions can be made from them.

¹ **CharingCross Hospital, London W6 8RF**

J POWELL-TUCK, MD, MRCP, senior medical registrar, gastroenterology unit

K D MacRAE, PHD, FIS, reader in medical statistics, department of medicine

R A PARKINS, MD, FRCP, consultant physician, gastroenterology unit

Division of Medical Statistics and Epidemiology", London School of Hygiene and Tropical Medicine, London WC1E 7HT

M J R HEALY, ma, professor of medical statistics

Gastroenterology" Unit, St Mark's Hospital, London EC1V 2PS

J E LENNARD-JONES, MD, MRCP, professor of gastroenterology

Correspondence to: Dr. Powell-Tuck.

Trial 1: Is nifedipine effective for the pain of the irritable bowel syndrome?

Nifedipine, a calcium channel blocking agent, reduces the tone of the lower oesophageal sphincter. Given sublingually it is absorbed rapidly with peak concentrations in the blood being achieved at 10 minutes and a pharmacological effect on smooth muscle lasting about 40 minutes,' Recurrent abdominal pain probably related to contraction of intestinal circular muscle and increased intraluminal pressure is a troublesome feature of the irritable bowel syndrome,' and a recent abstract suggested that nifedipine diminishes the abnormal colonic motor response of the rectosigmoid to distention in this disease.⁶ Patients need a rapidly effective remedy which can be used as required and which does not commit them to taking drugs regularly. We therefore conducted a controlled trial of sublingual nifedipine used for the pain of the irritable bowel syndrome. From the data on only 13 patients it was clear that in the dose chosen nifedipine was not usefully superior to placebo and that a large trial was unnecessary.

PATIENTS, METHODS, AND RESULTS

Six normal subjects tried sublingually and in randomised sequence one each of the active (nifedipine 5 mg) and placebo capsules to be used in the study and were asked to identify the active drug. Three were correct and three incorrect.

Thirteen patients with the irritable bowel syndrome who gave informed consent were included in the study. All had recurrent abdominal pain with distension and irregular bowel habit. In all patients results of guaiac stool testing with Hemocult and Fecatwin on three consecutive days each were negative and a full blood count and serum transaminase and alkaline phosphatase estimations were normal. None had any relevant abnormal finding on full examination including sigmoidoscopy. Where appropriate further investigation such as microscopy of duodenal juice or stool specimens, endoscopy, barium studies, or ultrasonography had been performed and showed nothing abnormal. None had responded adequately to a high fibre diet and antispasmodic agent the high fibre diet was maintained throughout. Each patient was supplied with three pairs of identical bottles labeled A and B, C and D, and E and F. One bottle of each pair contained active nifedipine (5 mg capsules), the other placebo. Patients tested the contents of bottle A as a buccally absorbed treatment for a bout of moderate to severe pain. If the first capsule was not effective within 20 minutes they took the second capsule from the same bottle. The result was recorded together with any side effects on a diary card. The contents of bottle B were tried on another day for a similar attack of pain, the result again being recorded and a preference, if any, given for either treatment. The contents of bottles C and D and E and F were similarly compared in pairs in later, similar episodes of abdominal pain if these occurred. Thus there was a maximum of three paired comparisons per patient.

Table I shows the results. All 13 patients made useful comparisons of the treatments, two registering an overall preference for the active treatment, six an overall preference for the placebo, and five no preference. In the 34 paired comparisons actually made 20 preferences were given, of which seven favoured the active drug and 13 the placebo. On 14 occasions the treatments were considered to be equal. Side effects recorded could be divided into mild faintness (placebo four cases, active drug one), flushing (placebo none, active one), nausea (placebo two, active one), and headache (placebo none, active five). Headache is the most

commonly recorded symptomatic side effect of nifedipine.⁷

TABLE I—*Trial 1. Preferences given in each paired comparison of nifedipine versus placebo and overall preference from combination of each patient's comparisons*

Case No	Age (yrs)	Sex	Preference*			Overall preference
			1	2	3	
1	20	M	A	A	O	A
2	29	F	P	A	O	O
3	41	F	P	O	P	P
4	50	F	P	P	A	P
5	25	M	O	-	-	O
6	26	F	P	A	-	O
7	63	F	O	O	A	A
8	25	F	P	O	P	P
9	46	F	O	O	O	O
10	20	F	O	P	A	O
11	31	M	P	O	O	P
12	52	F	P	O	P	P
13	57	F	-	P	-	P

*A=Nifedipine P= Placebo. O= No preference - = Inadequate or no comparison

COMMENT

Nifedipine was not significantly superior to placebo. What is the chance that in reality and in a much larger study nifedipine might against the trend of this study, prove to be useful?

The response of pain to treatment may be considered in two ways. Firstly, the individual overall preferences of the patients may be compared. If the two treatments were in reality equal we should expect a 50% preference for each in a large study. In fact, we saw only a 25% (2/8) preference for nifedipine with 95% confidence limits of 7 0-59 7%. Thus at best nifedipine is unlikely to be preferred by more than 60% of patients with 40% preferring placebo.

Secondly, the two preparations may be compared by assuming that all the patients with irritable

bowel syndrome would respond to an active drug similarly in which case all the paired comparisons can be pooled. In this case, of the 34 comparisons, seven favoured nifedipine and 13 favoured placebo - a preference for nifedipine of only 35% with 95% confidence limits of 17-57%. In this case there would be at best a 57% preference for nifedipine compared with a 43% preference for placebo. Among the 13 patients five developed headache after taking the active treatment, while none had this symptom after taking placebo. The incidence of this well known side effect together with the very unpromising response rate to nifedipine suggests that this small study needs no further extension and eliminates nifedipine used in this way as clinically useful for relieving the pain of the irritable bowel syndrome.

Using confidence limits is an important key to the analysis of small trials. In clinical practice there are often factors such as expense, convenience, and pharmaceutical stability which have to be considered when choosing between treatments. If a small trial excludes large clinical advantages for the more expensive, less convenient, or less stable preparation clinicians will choose the preparation which is cheap, convenient and stable. This is illustrated in the next trial.

Trial 2: Comparison of enemas containing 1 g and 2 g S-aminosalicylic acid in ulcerative proctosigmoiditis

Enemas containing 0.7 g (8) and 4.0 g (9) 5-aminosalicylic acid (mesalazine) have been shown to be effective in distal ulcerative colitis. More than 1 g/dl water is difficult to maintain in suspension, and it is difficult to prevent oxidation of larger amounts. The cost of purified mesalazine would represent a high proportion of the cost of making such enemas available commercially, and there is a theoretical risk of renal damage with high dose mesalazine absorbed systemically. Advice from a commercial company interested in marketing mesalazine enemas suggested that 1 g enemas would be easy and cheap to market and that 2 g enemas would be the highest dose which could be marketed practically, given the cost and expected shelf life. A double blind controlled trial of 1 g versus 2 g enemas was therefore conducted in patients with ulcerative proctosigmoiditis.

PATIENTS, METHODS, AND RESULTS

Twenty five patients with ulcerative colitis diagnosed clinically, sigmoidoscopically and histologically and shown by barium enema examination to be confined to the rectosigmoid agreed to a double blind controlled trial of enemas containing 1 g and 2 g mesalazine/dl water. Patients were included if they had active disease and were taking no medication for it other than sulphasalazine, which was continued unaltered. Twelve patients (four men; mean age 49, five taking sulphasalazine) were randomised to receive a 1 g retention enema each night and 13 (seven men; mean age 45; taking sulphasalazine) to receive a 2 g enema, and five to receive a 2g enema. Four patients randomised to receive 1 g enema and five to receive 2g enemas were in their first year of symptomatic disease. The trial lasted 28 days with assessment at entry (day zero), day 14, and day 28. Assessment was by clinical grading of malaise, bowel frequency, stool consistency, rectal bleeding, sigmoidoscopy, and rectal biopsy and by standard questions to determine whether there was adequate retention of the enemas and whether there was side effects. Results of clinical, sigmoidoscopic, and histological assessments were graded 0, 1 and 2, with 2 being the most severe. Bowel frequency was expressed as the number of bowel actions in the 24 hours preceding assessment.

The two groups had similar mean scores for the variables assessed at day zero except that there was a significantly higher ($p < 0.05$) score for rectal bleeding in the patients randomised to receive 2 g enemas. Table II shows the numbers of patients graded 0 for each variable at days 14 and 28. A patient was said to be in remission if the score was 0 for all the clinical variables and sigmoidoscopy showed a non-friable e rectal mucosa (graded 0). One patient developed worsening diarrhoea with the 2 g enemas and was withdrawn at day 5. No other side effects were noted and all patients retained the enemas satisfactorily overnight.

TABLE II—*Trial 2. Number of patients graded 0 in each group at days 14 and 28 for each variable with difference in percentage graded 0 between groups and 95% confidence limits for that difference.*

	Day 14				Day 28			
	No graded 0		Difference in % graded 0	95% Confidence limits of difference in % graded 0	No graded 0		Difference in % graded 0	95% Confidence limits of difference in % graded 0
	1g	2g			1g	2g		
Frequency	7	6	12	-27 to 51	9	7	28	-10 to 66
Bleeding	11	4	61	23 to 99	10	6	45	7 to 83
Malaise	8	9	-3	-39 to 34	10	11	6	-20 to 33
Stool Consistency	10	7	30	-7 to 66	10	10	14	-16 to 44
Sigmoidoscopy	7	4	39	-2 to 80	9	6	36	-3 to 75
Histology	7	3	33	-6 to 73	9	4	51	11 to 91
Remission	3	2	15	-19 to 49	7	4	33	-7 to 73

- At day 14 n=10 for sigmoidoscopy and remission as two patients declined sigmoidoscopy; at day 28 n=11 for 1g group (one exclusion at day 14).
- *Bowel frequency two or less daily
- Remission defined in text

COMMENT

The trial showed no tendency for 2 g enemas to be more effective than 1 g enemas. There was significantly less rectal bleeding at days 14 and 28 in patients given the 1 g enemas ($p < 0.01$, $p < 0.01$) but this was also true at day zero ($p < 0.05$) as judged by Wilcoxon rank sum testing.

Nevertheless, if anything there tended to be a greater improvement in rectal bleeding in the 1 g group.

The tables of Fleiss show that to have a 95% likelihood of finding a difference in hypothetical true response rates of 20% (for example, 65% and 85%) at $p < 0.05$ would require a trial with 138 patients in each group.¹² Can useful conclusions therefore be drawn from this trial containing 25 patients in all? Table II shows that at day 14 only malaise showed an overall advantage—that is, a negative difference in the percentage of patients graded 0—for the 2 g enemas, and this by only 3%. The 95% confidence limits show that, although not much confidence can be placed at day 14 in the assessment of frequency and malaise, advantages in favour of the 2 g enema of more than 7%, 6%, and 2% as judged by stool consistency, sigmoidoscopy, or histological appearances may be ruled out with 95% confidence. Similar results were obtained at day 28 when, as judged by four of the six assessment variables, an advantage of more than 10% in favour of the 2 g group may be ruled out with 95% certainty. It appears unlikely that 2 g enemas would produce remission at day 28 more than 7% more often than 1 g enemas. From both the marketing and clinical points of view this trial therefore provides enough evidence that the cheaper, more stable, and potentially safer 1 g enemas are not sufficiently less effective than 2 g enemas to warrant use of the latter and that the lower dose should be made available for routine use. Continuing the trial to include over 275 patients would be unlikely to alter this practical decision, would result in the problems of a multicentre trial, would delay practical action, and would be wasteful of financial and personnel resources.

Trial 3: Are prednisolone enemas the corticosteroid treatment of first choice in distal ulcerative colitis?

A previous study showed that active prednisolone appears in the blood after administration of prednisolone 21-phosphate rectally, in concentrations of the same order as those seen after an equivalent dose of oral prednisolone.¹³ Hence at least some and possibly much of the known therapeutic effect of these enemas might be mediated systemically. If a topical effect is also important it is likely that a substantial dose of prednisolone given rectally would be noticeably more effective than the same dose given by mouth. A trial comparing equivalent doses of rectal prednisolone phosphate and oral prednisolone was therefore conducted in patients with active distal ulcerative colitis who would normally have been treated with prednisolone phosphate enemas.

PATIENTS, METHODS AND RESULTS

We studied 36 patients with ulcerative colitis diagnosed clinically, sigmoidoscopically, and histologically and shown by barium enema examination not to extend proximal to the splenic flexure. The study was conducted as a single blind controlled trial comparing an oral dose of prednisolone 20 mg given once each morning with an equivalent dose of prednisolone phosphate dissolved in tap water and administered as a retention enema each night. Clinicians, unaware of which treatment had been allocated, assessed the patients at entry into the trial and at days 14 and 28. The parameters used for assessment were malaise and the presence of abdominal tenderness graded 0-3; abdominal pain, bowel frequency, Stool consistency, and the presence of extraintestinal complications of ulcerative colitis graded 0-2; and the presence of anorexia or of nausea or vomiting graded 0-1. Sigmoidoscopic appearances and histological assessments of rectal mucosal biopsy samples were each graded 0-2. Sixteen patients (seven men; mean age of

46 (6; 10 receiving sulphasalazine) were treated by mouth and 20 (eight men; mean age 43.4; 13 receiving sulphasalazine) were treated rectally. Only one patient (treated rectally) was in his first attack of colitis. Table III gives the overall results. No individual measures of activity clearly responded differently to the two treatments.

In addition to outcome, the following side effects of corticosteroids were assessed: change in appetite and mood, dyspepsia, hirsutism, acne, striae, oral candidiasis, peripheral oedema, hypertension, hypokalaemia, and glycosuria. Only one statistically significant difference was observed between the two groups: of 14 patients in the oral group and 19 in the rectal group who had regular blood pressure measurements, 3 (21%) and 13 (68%), respectively, showed a rise in diastolic pressure of 5 mm Hg or more ($p < 0.01$). When a less strict definition of raised blood pressure was applied (rise in diastolic pressure > 10 mm Hg) it was seen in 1 (7%) and 2 (11%) of the patients, respectively. Six of 16 patients treated by mouth and four of 19 patients treated rectally had no observed side effects during the 28 days of the trial.

COMMENT

Treatment of distal ulcerative colitis by retention enemas is widely accepted as the treatment of first choice. Nevertheless, aside from their obvious inconvenience for the patient, they have several practical therapeutic disadvantages. They tend to be given at night when absorbed corticosteroids cause maximal adrenal suppression. The inflamed rectum is less able than normal to tolerate any introduced volume and patients with severe distal colitis sometimes cannot retain a therapeutic enema." No controlled data support the widespread clinical impression that prednisolone phosphate enemas produce fewer side effects than equivalent doses given by mouth and, by virtue of a local effect, are more effective than oral administration. This study lent no support to the contention that rectal steroids produce fewer clinically observable side effects than a similar dose of oral prednisolone given once each morning over 28 days.

The overall clinical results may be interpreted as showing a small tendency in favour of the rectal group at day 14 which was no longer apparent by day 28. In order to consider the relative efficacy of the two treatments we can first estimate the percentage of patients who would respond to enema treatment in a larger study. Table IV shows the results of this study combined with those of two others, which were also conducted at, or closely associated with, St Mark's Hospital and in which rectal prednisolone phosphate 20 mg was assessed under very similar controlled conditions." " Seventy eight per cent responded by day 14 and 82% by day 28. The tables of Fleiss show that to have a 95% chance of detecting a true response rate to oral treatment 20% less than the 80% seen for rectal treatment at a significance level of 0.05 we should need 153 patients in each group.' Similarly, to show a true 10% difference would need 524 patients in each group. It was therefore thought impracticable to continue the trial. Nevertheless, some important conclusions can be drawn from the data available. Table V simplifies the results.

TABLE IV -- Trial 3. Results of enema treatment with prednisone 21-phosphate of equivalent dosage to 20mg prednisolone

	Day 14				Day 28			
	Remission	Improved	No Change	Worse	Remission	Improved	No Change	Worse
Oral	3	8	4	0	7	6	2	1
Rectal	6	7	5	1	7	5	5	0

*One patient in oral group seen only at day 28.

*In rectal group one patient did not return after entry, one was withdrawn at day 14 because she could retain enemas, and another did not return at day 28.

Remission=Symptoms, signs, and sigmoidoscopy score 0. Improved=Reduction from score on day zero by 2 or more. No change=Change from day zero score<2. Worse=Increase from day zero score by 2 or more.

At day 14 the difference in the proportions of patients responding to the two treatments was 5% with 95% confidence limits of +35.8 to -25.8. At day 28 the difference was 109% with 95% confidence limits of +39.2 to -19.2. Hence at day 14 we have essentially excluded an advantage of enemas over tablets of more than 26% and at day 28 of more than 19%. (The same results may be obtained by calculating the 95% confidence limits of the difference between the odds ratios of response and failure to respond to treatment for the two treatments.²⁰) Applying these results to our best estimates of real percentage response rates to enemas at days 14 and 28 (table IV) shows that the true response rates in a large series to oral prednisolone given in this way are unlikely to be worse than 52% (day 14) and 62% (day 28) and may well be better. Side effects seemed to be no more pronounced over 28 days with oral rather than rectal treatment. It might therefore be reasonable to hold the inconvenient and unpleasant enema treatment in reserve for those patients whose distal colitis does not respond to a course of 20 mg oral prednisolone given over 14-28 days.

TABLE V-Trial 3. Simplified results

	DAY 14		DAY 28	
	ORAL	RECTAL	ORAL	RECTAL
No of patients treated	15	19	16	17
No (%) improved	11(73)	13(68)	13(81)	12(71)

Discussion

Brown has lucidly described the problems of calculating the sample size for a clinical trial and comments particularly on the misconception that calculation of sample size is an objective and

unexceptionable procedure.²¹ Rather, sample size calculations are approximations which depend on assumptions and arbitrary judgments. Brown comments that it is often wise to approach this problem by determining the resources available and then calculating the discriminating power that can be achieved. This was the approach that we used, studying patients fulfilling strict inclusion criteria who could be recruited in a specified time.

An alternative approach is to use a sequential design. Sequential designs are best suited to studies which have a single all important criterion of success or failure which becomes evident quickly after treatment; they are less suitable when multiple end points are relevant and also have the practical disadvantage that the duration of the study is not known in advance.

Large trials may arguably be more representative than small trials, but this argument assumes that—for example, in a multicentre trial the effect of treatment differs importantly from centre to centre. Often this assumption is false, but even if true it is frequency not adequately taken into account in the analysis. Disadvantages of large trials include their expense, the length of time before a clinical decision can be made, and the enormous task of administration and coordination.

The large trial has an important but restricted role in clinical research namely, the detection of small differences in important outcomes, such as survival, in common diseases. Much of clinical medicine, however, is concerned with relief of symptoms in settings where a treatment is worth while only if it produces a large benefit relative either to its side effects or to other remedies which are cheaper or more convenient.

The most important aspect of the interpretation of any trial, large or small, is to consider the confidence interval for the difference that is likely to exist. Overconcentration on the simplistic notion of statistical significance is responsible for most of the ill based criticisms of small trials. Hence we, like Blackwelder,²² emphasise the use of confidence intervals in the interpretation of trial data; their role in judging the results of small trials is clearly shown in the three examples given in this paper.

Schwartz and Lellouch distinguished between "explanatory" and "pragmatic" approaches to clinical trials.²³ With the explanatory approach the main objective is to further scientific knowledge, while the pragmatic approach aims to enable a practical decision to be made on which two treatments should be used. In pragmatic terms the error to be avoided in clinical practice after a study is opting for the worse of compared treatments—what Schwartz and Lellouch termed a type III error. Instead, so called explanatory trials concentrate on minimising the type I and type II errors. A type I error is incorrectly preferring one of two equivalent treatments, an error which in clinical terms is unimportant. A type II error made by concluding that two treatments are equivalent is easily avoided by always selecting the apparently superior treatment regardless of statistical evidence. The real worry clinically is the type III error, in which a clinically significantly inferior treatment is preferred to a superior one on the basis of insufficient data. Schwartz and Lellouch argued that most trials wrongly take the explanatory approach and thus necessitate conventional significance testing of the null hypothesis. Often the clinician seeks to answer practical questions; he wants to know not whether a new treatment is 'effective', but whether it is "effective enough" to oust previous treatment. Under certain circumstances all that is needed is for a trial to show that there is or is not a large difference in efficacy between two treatments.

These three trials are presented as examples of the pragmatic approach to clinical trials. They are offered to show that practical decisions can often be made in clinical medicine without reference to the null hypothesis and that the key to analysis lies more with a critical use of confidence limits than with conventional significance testing. They show how a small trial can disclose a statistically insignificant risk of making a clinically important type III error. We hope that they will stimulate the conduct and more thoughtful analysis of small studies and discourage journal referees from rejecting trials solely because they are "too small."

We thank Dr M J Glynn for help with word processing during the preparation of this paper. Dr S Shousha (Sharing Cross Hospital) made the histological assessments in trial 2 and Dr B C Morson (St Mark's Hospital) made them in trial 3. We are grateful to Dr G Macdonald, of Bayer Ltd, and Dr B Donovan and Mr M Tupholme, of Ferring Pharmaceuticals, for providing drugs and placebos for trials I and 2, respectively.

References

1. Freiman JA, Thomas AB, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomised controlled trial. *N Engl J Med* 1978;299:690-4
 2. Blackwell JN, Holt S, Heading RC. Effect of nifedipine on oesophageal motility and gastric emptying. *Digestion* 1981;21:50-60.
 3. Foster TA, Hamann SR, Richards VR, et al. Nifedipine, kinetics and bioavailability after single intravenous and oral doses in normal subjects. *J Clin Pharmacol* 1983;23:161-70.
 4. Bartolotti M, Labo GI. Clinical and manometric effects of nifedipine in patients with esophageal achalasia. *Gastroenterology* 1981;80:39-44.
 5. Burns TW. Colonic motility in the irritable bowel syndrome. *Arch Intern Med* 1980;140:247-51.
 6. Blume M, Schuster M, Tucker H. Effect of nifedipine on colonic motility on the irritable bowel syndrome. *Gastroenterology* 1983;84:1109.
 7. Leonard RG, Talbert RL. Calcium channel blocking agents. *Clinical pharmacy* 1982;1:17-33.
 8. Azad Khan KA, Piris J, Truelove SC. An experiment to determine the active therapeutic moiety of suphasalazine. *Lancet* 1977;ii:892-5.
 9. Campieri M, Lanfranchi GA, Bazzocchi G, et al. Treatment of ulcerative colitis with high dose 5 aminosalicylic acid enemas. *Lancet* 1981;ii:270.
 10. Calder IC, Fincher CC, Gran CR, Ham KN, Tange JD. Nephrotoxic lesions from 5 aminosalicylic acid. *Br Med J* 1972;I:152-4.
 11. Powell-Tuck J, Day DW, Bucknell NA, Wadsworth J, Lennard-Jones JE. Correlations between defined sigmoidoscopic appearances and other measures of disease activity ulcerative colitis. *Dig Dis Sci* 1982;27:533-7.
 12. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York:Wiley, 1981:260-80.
 13. Powell-Tuck J, Lennard-Jones JE, May CS, Wilson CE, Paterson JW. Plasma prednisolone levels after administration of prednisolone 21-phosphate as a retention enema in colitis. *Br Med J* 1976;I:193-5.
 14. Sparberg M, Jensen R, Beering SC, Shannon I. Pituitary adrenal axis suppression following intrarectal administration of dexamethasone 21-phosphate. *Gastroenterology* 1967;52:519-20.
 15. Nichols T, Nugent CA, Tyler FH. Diurnal variation in suppression of adrenal function by glucocorticoids. *Journal of Clinical Endocrinology* 1965;25:343-9.
 16. Myles AB, Bacon PA, Daly JR. Single daily dose corticosteroid treatment: effect of adrenal function and therapeutic efficacy in various disease. *Ann Rheum Dis* 1971;30:149-53.
 17. Farthing MJG, Lennard-Jones JE. Sensibility of the rectum to distension and the anorectal distension reflex in ulcerative colitis. *Gut* 1978;19:64-9.
 18. Misiewicz JJ, Connell AM, Lennard-Jones JS, Jones FA. Comparison of oral and rectal steroids in the treatment of proctocolitis. *Proceedings of the Royal Society of Medicine* 1964;57:561-2.
 19. Multicentre Trial. Betamethasone 17 valerate and prednisolone 21-phosphate retention enemata in proctocolitis. *Br Med J* 1971;iii:84-6.
 20. Armitage P. Statistical methods in medical research. Oxford:Blackwell, 1971.
 21. Brown BW. Statistical controversies in the design of clinical trials-some personal views. *Controlled Clin Trials* 1980;1:13-27.
 22. Blackwelder WC. "Proving the null hypothesis" in clinical trials. *Controlled Clin Trials* 1982;3:345-53.
 23. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;20:637-48.
- (Accepted 26 November 1985)