Statistics Notes
Analysis of a trial randomised in clusters

Sally M Kerry, J Martin Bland

Division of General Practice and Primary Care, St George's Hospital Medical School, London SW17 0RE
Sally M Kerry, *lecturer in medical statistics*

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
J Martin Bland, *professor of medical statistics*

Correspondence to: Mrs Kerry.

A cluster randomised study is one where a group of subjects are randomised to the same treatment together—for example, when women in some districts are offered breast cancer screening and compared with women in other districts, or when the patients of general practitioners who have been given special training are compared with the patients of those who have not.[1]

Several techniques exist for analysing the data from such studies, but the essence of them is that the experimental unit (district or general practitioner) is the unit of analysis.[2] A simple approach is to construct a summary statistic for each cluster and then analyse these summary values. The idea is similar to the analysis of repeated measurements on the same subject, where we construct a single summary statistic over the times for each individual.[3] The same principle is used in meta-analysis.

Here we shall describe the analysis of such a study, an investigation of the effect of guidelines for radiological referral on the referral practice of general practitioners.[4] Thirty four practices referring patients to St George's Hospital for *x* ray examinations were randomised into two groups. The practices in the intervention group received a one page laminated copy of extracts from the Royal College of Radiologists guidelines for radiological referral adapted for general practitioners and a covering letter. The control group practices were not sent anything. Thus all patients in a practice were subject to the same intervention and the practice was the experimental unit. The outcome measure was the percentage of *x* ray examinations requested that conformed to the guidelines (table 1).

The mean percentages in the two groups can be compared by the two sample *t* method. The observed difference is $81.6 - 73.6 = 8.0$ and the standard error of the difference is 4.3. There are 32 degrees of freedom, giving a 95% confidence interval of $8.0 \pm 2.037 \times 4.3$, or -1 to 17 percentage points. For the test of significance, the test statistic is $8.0/4.3 = 1.86$. This gives $P = 0.07$. If the assumptions for the *t* test did not hold the data might be transformed[5] or the Mann Whitney U test could be used.

The standard two sample *t* method gives equal weight to all practices, despite the widely varying numbers of referrals. It is preferable to carry out a weighted analysis, using the numbers of referrals as the weights. This gives the estimated difference in mean percentage of appropriate referrals as 7.0 percentage points, standard error 3.3, $t = -2.10$, 95% confidence interval 0.2 to 13.8, $P = 0.04$. We shall give the details of this approach in a separate Statistics Note.

If we were to act as if we had randomly assigned individual patients to intervention groups then we would calculate the difference in the proportion of requests in each group conforming to the guidelines. We could then use the usual normal approximation to test the observed difference between the proportions in the two groups. There were a total of 429 requests from 17 practices in the intervention group, of which 79.5% conformed with the guidelines. In the control group 72.5% of 702 requests from 17 practices conformed to the guidelines. The difference in percentage conforming is 7.0, and the standard error of the difference is 2.6, giving a 95% confidence interval of 2 to 12 percentage points, $P = 0.008$ (χ^2 test). Ignoring the clustering results in confidence intervals which are too narrow and P values which are too small; hence it is likely to produce spuriously significant differences.

As there will be some loss of power owing to the variability between clusters, sample size calculations need to take this into account and consider not only the number of patients but the number of clusters as well. We shall discuss this topic in a future Statistics Note.

1 Bland JM, Kerry SM. Statistics Notes. Trials randomised in clusters. *BMJ* 1997;315:600.
2 Altman DG, Bland JM. Statistics Notes. Units of analysis. *BMJ* 1997;314:1874.
3 Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990;300:230-5.
4 Oakeshott P, Kerry SM, Williams JE. Randomised controlled trial of the effect of the Royal College of Radiologists' guidelines on general practitioners' referral for radiographic examination. *Br J Gen Pract* 1994;44:197-200.
5 Bland JM, Altman DG. Statistics Notes. The use of transformations when comparing two means. *BMJ* 1996;312:1153.

Table 1 Number of requests conforming to guidelines for each practice in the intervention and control groups

	Intervention group			Control group		
	No of requests			No of requests		
Practices	Total	Conforming	% Conforming	Total	Conforming	% Conforming
1	20	20	100	7	7	100
2	7	7	100	37	33	89
3	16	15	94	38	32	84
4	31	28	90	28	23	82
5	20	18	90	20	16	80
6	24	21	88	19	15	79
7	7	6	86	9	7	78
8	6	5	83	25	19	76
9	30	25	83	120	90	75
10	66	53	80	88	64	73
11	5	4	80	22	15	68
12	43	33	77	76	52	68
13	43	32	74	21	14	67
14	23	16	70	126	83	66
15	64	44	69	22	14	64
16	6	4	67	34	21	62
17	18	10	56	10	4	40
Total	429	341		702	509	
Mean (SD)			81.6 (11.9)			73.6 (13.1)