3 Posner T, Vessey M. *Prevention of cervical cancer: the patient's view*. London: King Edward's Fund for London, 1988.
4 Nathoo V. Investigation of non-responders at a cervical cancer screening clinic in Manchester. *BMJ* 1988;296:1041-2.
5 Sansom CD, MacInerney J, Oliver V, Wakefield J. Differential response to recall in a cervical screening programme. *British Journal of Preventive and Social Medicine* 1975;29:40-7.
6 Bailie R, Petrie K. Women's attitudes to cervical smear testing. *N Z Med J* 1990;103:292-5.
7 Davison RL, Clements JE. Why don't they attend for a cytotest? A pilot study among a high-risk population. *The Medical Officer* 1971;1:329-31.
8 Cullum DE, Savory JN. Patient preference for cervical cytology. *BMJ* 1983;287:329-32.
9 Pierce M, Lundy S, Palanisamy A, Winning S, King J. Prospective randomised controlled trial of methods of call and recall for cervical cytology screening. *BMJ* 1989;299:160-2.
10 Elkind AK, Haran D, Eardley A, Spencer B. Reasons for non-attendance for computer-managed cervical screening: pilot interviews. *Soc Sci Med* 1988; 27:651-60.
11 McKie L, Gregory S, Garrett A, Bellerby I. *Patterns of take up and attitudes towards smear tests in the County of Cleveland (1989/1990)*. Stockton-on-Tees: Community Health Council (North Tees), 1990.
12 Bowling A, Jacobson B. Screening: the inadequacy of population registers. *BMJ* 1989;298:545-6.
13 Allman ST, Chamberlain J, Harman P. The national cervical cytology recall scheme: a report of a pilot study. *Health Trends* 1974;6:39-41.
14 Beardow R, Oerton J, Victor C. Evaluation of the cervical cytology screening programme in an inner city health district. *BMJ* 1989;299:98-100.
15 Ross SK. Cervical cytology screening and government policy. *BMJ* 1989;299: 101-4.
16 Shroff KJ, Corrigan AM, Bosher M, Edmonds MP, Sacks D, Coleman DV. Cervical screening in an inner city area: response to a call system in general practice. *BMJ* 1988;297:1317-8.
17 Kitzinger J. The methodology of focus groups: the importance of interaction between research participants. *Sociology of Health and Illness* 1994:16: 103-21.
18 Naish J. *General practitioners' view of cervical cytology screening in City and east London*. London: City and East London FHSA, 1992.

(Accepted 16 September 1994)

Statistics Notes

Matching

J Martin Bland, Douglas G Altman

This is the ninth in a series of occasional notes on medical statistics

In many medical studies a group of cases, people with a disease under investigation, are compared with a group of controls, people who do not have the disease but who are thought to be comparable in other respects. This happens in epidemiological case-control studies, where a possible risk factor is compared between cases and controls to investigate the cause of the disease, and in clinical studies, where the characteristics of cases and controls are compared to investigate the nature of the disease. In both types of study cases and controls are sometimes matched. This means that for every case there is a control who has the same (or closely similar) values of the matching variables. Matching may be by sex, age to within five years, ethnic group, etc. Sometimes there are two or more such controls for each case.

We match to ensure that controls and cases are similar in variables which may be related to the variable which we are studying but are not of interest in themselves. For example, in many epidemiological case-control studies age is an important predictor of exposure to the risk factor under investigation. There are strong cohort effects in variables such as cigarette smoking and diet. If we do not take age into account we may get spurious differences between cases and controls because, for example, cases are older than controls. Matching ensures that any difference between cases and controls cannot be a result of differences in the matching variables. However, we cannot then examine the effects of the matching variables.

Sometimes matching is ignored in the analysis of the data. If the matching variables are important, this is inefficient. Matching variables, such as age and sex, may be strongly related to the variable of interest. If we allow for the matching in the analysis the variation due to these variables is removed. If we ignore the matching the variability which is related to the matching variables becomes part of the unexplained variation and may obscure important differences. For example, if we compare the mean blood pressure of subjects with a disease to that of their age matched controls, the variability in blood pressure which is associated with its increase with age will be part of the residual variance and will increase the standard error of the difference between the means. Instead, we should use the differences between individually matched cases and their controls. Appropriate simple methods include the paired t test for means, McNemar's test for proportions, and the sign test for ordinal data. Sometimes there is no suitable method of matched analysis, as in survival analysis. We can usually adjust for the matching variables, however.

It is desirable to adjust for matching when this was done to make the groups comparable for believed prognostic or confounding variables. This should be done even if in the sample the variable is not significantly prognostic or confounding. By contrast, matching is sometimes merely a convenient method of drawing the sample. For example, in studying cot deaths we might take as a control the next birth in the same hospital. This is sometimes referred to as cosmetic matching. We can ignore the matching in the analysis of such studies.

There are disadvantages to matching. If we match we can only use cases for whom we have matching controls. The more variables we match on the more difficult it may be to find such controls. Even to match on age, sex, and ethnic group we need a large population of potential controls from which to draw. A practical difficulty with matched pairs is that if we want to adjust for other, non-matched, variables the analysis required is more complex than ordinary multiple or logistic regression.

In a large study with many variables it is easier to take an unmatched control group and adjust in the analysis for the variables on which we would have matched, using ordinary regression methods. Matching is particularly useful in small studies, where we might not have sufficient subjects to adjust for several variables at once.

Some authors use "matched" to mean that the two groups are similar in the distribution of the matching variables, but not that there is individual matching of each case to his or her own control. Such studies should not be described as matched.

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

Imperial Cancer Research Fund, PO Box 123, London WC2A 3PX
Douglas G Altman, *head*

Correspondence to:
Mr Bland.

BMJ 1994;309:1128