4 Gibbins RL, Riley M, Brimble P. Effectiveness of programme for reducing cardiovascular risk for men in one general practice. *BMJ* 1993;**306**:1652-6.
5 Family Heart Study Group. Randomised controlled trial evaluating cardiovascular screening and intervention in general practice: principal results of British family heart study. *BMJ* 1994;**308**:313-20.
6 Imperial Cancer Research Fund OXCHECK Study Group. Effectiveness of health checks conducted by nurses in primary care: results of OXCHECK study after one year. *BMJ* 1994;**308**:308-12.
7 Silagy C, Muir J, Coulter A, Thorogood M, Roe L. Cardiovascular risk and attitudes to lifestyle: what do patients think? *BMJ* 1993;**306**:1657-60.
8 Rosenbrg L, Kaufman DW, Helmrich SP, Shapiro S. The risk of myocardial infarction after quitting smoking in men under 55 years of age. *N Engl J Med* 1985;**313**:1511-4.
9 Todd I, Ballantyre D. Effect of exercise training on the total ischaemic burden: an assessment by 24 hour ambulatory electrocardiographic monitoring. *Br Heart J* 1992;**68**:560-6.
10 Law MR, Wald NJ, Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? *BMJ* 1994;**308**:367-73.
11 Ornish D, Brown SE, Scherwitz LW, Billings JH, Armstrong WT, Ports TA. Can lifestyle changes reverse coronary heart disease? The lifestyle heart trial. *Lancet* 1990;**336**:129-33.
12 *Change of heart. Strategy of prevention of coronary heart disease in Northern Ireland.* Belfast: DHSS, 1986.
13 Shaper AG, Pocock SJ, Phillips AN, Walker M. Identifying men at high risk of heart attacks: strategy for use in general practice. *BMJ* 1986;**293**:474-9.
14 Rose G, Hamilton PJS, Keen H, Reid DD, McCartney P, Jarret RJ. Myocardial ischaemia risk factors and death from coronary heart disease. *Lancet* 1977;i:105-9.
15 Daly LE, Hickey N, Mulcahy R. Course of angina pectoris after an acute coronary event. *BMJ* 1986;**293**:653-6.
16 Campbell MJ, Elwood PC, Abbas S, Waters WE. Chest pain in women: a study of prevalence and mortality follow up in South Wales. *J Epidemiol Community Health* 1984;**38**:17-20.
17 Neill WA, Branch LG, De Jong G, Smith NE, Hogan CE, Corcoran PJ, *et al.* Cardiac disability: the impact of coronary heart disease on patients' daily activities. *Arch Intern Med* 1985;**145**:1642-7.
18 Research Committee, Northern Region Faculty, Royal College of General Practitioners. Study of angina in patients aged 30 to 59 in general practice. *BMJ* 1982;**285**:1319-21.
19 Health Promotion Agency for Northern Ireland. *Research on nutrition in Northern Ireland.* Belfast: HPANI, 1992.
20 Medical Research Council Working Party. MRC trial of treatment of mild hypertension: principal results. *BMJ* 1985;**291**:97-104.
21 Dawber TR. *The Framingham study: the epidemiology of atherosclerotic disease.* Cambridge, MA: Harvard University Press, 1980:76-90.
22 Borkan GA, Sparrow D, Wisniewski C, Vokonos PS. Body weight and coronary disease risk: patterns of risk factor change associated with long-term weight change. *Am J Epidemiol* 1986;**124**:410-9.
23 Bingham SA. The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviws (series A)* 1987;**57**:720-1.
24 Thorgood M, Roe L, McPherson K, Mann J. Dietary intake and plasma lipid levels: lessons from a study of the diet of health conscious groups. *BMJ* 1990;**300**:1297-301.
25 Connor SL, Gustafson JR, Sexton G, Becker N, Artaud-Wild S, Connor WE. The diet habit survey: a new method of dietary assessment that relates to plasma cholesterol changes. *Journal of the American Diabetic Association* 1992;**92**:41-7.

(Accepted 24 August 1994)

Statistics Notes

Quartiles, quintiles, centiles, and other quantiles

Douglas G Altman, J Martin Bland

This is the eighth in a series of occasional notes on statistics

When presenting or analysing measurements of a continuous variable it is sometimes helpful to group subjects into several equal groups. For example, to create four equal groups we need the values that split the data such that 25% of the observations are in each group. The cut off points are called quartiles, and there are three of them (the middle one also being called the median). Likewise, we use two tertiles to split data into three groups, four quintiles to split them into five groups, and so on. The general term for such cut off points is quantiles; other values likely to be encountered are deciles, which split data into 10 parts, and centiles, which split the data into 100 parts (also called percentiles). Values such as quartiles can also be expressed as centiles; for example, the lowest quartile is also the 25th centile and the median is the 50th centile. We consider below some common applications of quantiles.

A common confusion is to use the terms tertiles, quartiles, quintiles, etc, not for the cut off points but for the groups so obtained, but these are properly called thirds, quarters, fifths, and so on.

Data description—The mean and standard deviation are useful to summarise a set of observations. When the data have a skewed distribution it is often preferable to quote instead the median and two outer centiles, such as the 10th and 90th. The first and third quartiles (25th and 75th centiles) are sometimes used; these define the interquartile range. The median is a useful summary statistic when some of the values are not actually measured—for example, because some values are outside the range of the measuring equipment. Similarly, the median is frequently used when summarising survival data, when it is usual for some of the survival times to be unknown.

Reference intervals and centiles—A special type of data description arises in the construction of a reference interval (normal range). A 95% reference interval is defined by the values that cut off 2½% at each end of the distribution. (These values are often quite reasonably called the 2½ and 97½th centiles, although it is not strictly correct to have half centiles.) Reference intervals are widely used in clinical chemistry. By contrast, charts for the assessment of human size or growth usually show several centiles.[1] Reference centiles are sometimes derived using the normal distribution,[2] in which case any new observation can be placed at a specific centile.

Analysis of continuous variables—Continuous variables, such as serum cholesterol concentration and lung function, are often categorised in statistical analyses. It is usual to use quantiles, so that there are the same number of individuals in each group. Such grouping discards information but may allow for simpler presentation, such as in tables. The fewer groups created the greater is the loss of information. In regression analyses continuous explanatory variables are often categorised into two or more groups. Although this slightly complicates the analysis, it avoids a direct assumption that there is a linear relation between the variable and the outcome of interest. However, it leads to a model in which risk apparently jumps at certain values of the predictor variable rather than increasing smoothly.

Calculation of quantiles—The calculation of centiles and other quantiles is not as simple as it might seem. The data should be ranked from 1 to n in order of increasing size. The kth centile is obtained by calculating $q=k(n+1)/100$ and then interpolating between the two values with ranks either side of the qth. For example, for the 5th centile of a sample of 145 observations we have $q=5\times146/100=7{\cdot}3$. We estimate the 5th centile as the value 0·3 of the way between the 7th and 8th ranked observations. If these data values are 11·4 and 14·9 the estimated centile is 12·45. Confidence intervals can be constructed for any quantile.[3]

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

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

Correspondence to:
Mr Altman.

BMJ 1994;**309**:996

1 Cole TJ. Do growth charts need a face lift? *BMJ* 1994;**308**:641-2.
2 Altman DG. *Practical statistics for medical research.* London: Chapman and Hall, 1991:419-26.
3 Campbell MJ, Gardner MJ. Calculating confidence intervals for some non parametric analyses. In: Gardner MJ, Altman DG,eds. *Statistics with confidence.* London: British Medical Journal, 1989:71-9.