

- 6 The Cardiac Arrhythmia Suppression Trial I Preliminary Report. Effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med* 1989;321:406-12.
- 7 Dickert N, Grady C. What's the price of a research subject? Approaches to payment for research participation. *N Engl J Med* 1999;341:198-203.
- 8 Grady C. Money for research participation: does it jeopardize informed consent? *Am J Bioethics* 2001;1:40-4.
- 9 Macklin R. "Due" and "undue" inducements: On paying money to research subjects. *IRB: a review of human subjects research* 1981;3:1-6.
- 10 McGee G. Subject to payment? *JAMA* 1997;278:199-200.
- 11 McNeil P. Paying people to participate in research. *Bioethics* 1997;11:390-6.
- 12 Wilkinson M, Moore A. Inducement in research. *Bioethics* 1997;11:373-89.
- 13 Halpern SD, Karlawish JHT, Casarett D, Berlin JA, Asch DA. Empirical assessment of whether moderate payments are undue or unjust inducements for participation in clinical trials. *Arch Intern Med* 2004;164:801-3.
- 14 Bentley JP, Thacker PG. The influence of risk and monetary payment on the research participation decision making process. *J Med Ethics* 2004;30:293-8.
- 15 Viens AM. Socio-economic status and inducement to participate. *Am J Bioethics* 2001;1.
- 16 Beauchamp TL, Childress JF. *Respect for autonomy. Principles of biomedical ethics*. 4th ed. New York: Oxford University Press, 1994:120-88.
- 17 Roberts LW. Evidence-based ethics and informed consent in mental illness research. *Arch Gen Psychiatry* 2000;57:540-2.
- 18 Bayer R, Oppenheimer GM. *Toward a more democratic medicine: sharing the burden of ignorance. AIDS Doctors: voices from the epidemic*. New York: Oxford University Press, 2000:156-69.
- 19 Coulter A, Rozansky D. Full engagement in health. *BMJ* 2004;329:1197-8.
- 20 Joffe S, Manocchia M, Weeks JC, Cleary PD. What do patients value in their hospital care? An empirical perspective on autonomy centred bioethics. *J Med Ethics* 2003;29:103-8.
- 21 Heesen C, Kasper J, Segal J, Kopke S, Muhlhauser I. Decisional role preferences, risk knowledge and information interests in patients with multiple sclerosis. *Mult Scler* 2004;10:643-50.
- 22 Azoulay E, Pochard F, Chevret S, et al. Half the family members of intensive care unit patients do not want to share in the decision-making process: a study in 78 French intensive care units. *Crit Care Med* 2004;32:1832-8.
- 23 Dunn LB, Gordon NE. Improving informed consent and enhancing recruitment for research by understanding economic behavior. *JAMA* 2005;293:609-12.

## Statistics Notes

### Standard deviations and standard errors

Douglas G Altman, J Martin Bland

The terms "standard error" and "standard deviation" are often confused.<sup>1</sup> The contrast between these two terms reflects the important distinction between data description and inference, one that all researchers should appreciate.

The standard deviation (often SD) is a measure of variability. When we calculate the standard deviation of a sample, we are using it as an estimate of the variability of the population from which the sample was drawn. For data with a normal distribution,<sup>2</sup> about 95% of individuals will have values within 2 standard deviations of the mean, the other 5% being equally scattered above and below these limits. Contrary to popular misconception, the standard deviation is a valid measure of variability regardless of the distribution. About 95% of observations of any distribution usually fall within the 2 standard deviation limits, though those outside may all be at one end. We may choose a different summary statistic, however, when data have a skewed distribution.<sup>3</sup>

When we calculate the sample mean we are usually interested not in the mean of this particular sample, but in the mean for individuals of this type—in statistical terms, of the population from which the sample comes. We usually collect data in order to generalise from them and so use the sample mean as an estimate of the mean for the whole population. Now the sample mean will vary from sample to sample; the way this variation occurs is described by the "sampling distribution" of the mean. We can estimate how much sample means will vary from the standard deviation of this sampling distribution, which we call the standard error (SE) of the estimate of the mean. As the standard error is a type of standard deviation, confusion is understandable. Another way of considering the standard error is as a measure of the precision of the sample mean.

The standard error of the sample mean depends on both the standard deviation and the sample size, by the simple relation  $SE = SD/\sqrt{\text{sample size}}$ . The standard error falls as the sample size increases, as the extent of chance variation is reduced—this idea underlies the sample size calculation for a controlled trial, for

example. By contrast the standard deviation will not tend to change as we increase the size of our sample.

So, if we want to say how widely scattered some measurements are, we use the standard deviation. If we want to indicate the uncertainty around the estimate of the mean measurement, we quote the standard error of the mean. The standard error is most useful as a means of calculating a confidence interval. For a large sample, a 95% confidence interval is obtained as the values  $1.96 \times SE$  either side of the mean. We will discuss confidence intervals in more detail in a subsequent Statistics Note. The standard error is also used to calculate P values in many circumstances.

The principle of a sampling distribution applies to other quantities that we may estimate from a sample, such as a proportion or regression coefficient, and to contrasts between two samples, such as a risk ratio or the difference between two means or proportions. All such quantities have uncertainty due to sampling variation, and for all such estimates a standard error can be calculated to indicate the degree of uncertainty.

In many publications a  $\pm$  sign is used to join the standard deviation (SD) or standard error (SE) to an observed mean—for example,  $69.4 \pm 9.3$  kg. That notation gives no indication whether the second figure is the standard deviation or the standard error (or indeed something else). A review of 88 articles published in 2002 found that 12 (14%) failed to identify which measure of dispersion was reported (and three failed to report any measure of variability).<sup>4</sup> The policy of the *BMJ* and many other journals is to remove  $\pm$  signs and request authors to indicate clearly whether the standard deviation or standard error is being quoted. All journals should follow this practice.

Competing interests: None declared.

- 1 Nagele P. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. *Br J Anaesthesiol* 2003;90:514-6.
- 2 Altman DG, Bland JM. The normal distribution. *BMJ* 1995;310:298.
- 3 Altman DG, Bland JM. Quartiles, quintiles, centiles, and other quantiles. *BMJ* 1994;309:996.
- 4 Olsen CH. Review of the use of statistics in *Infection and Immunity*. *Infect Immun* 2003;71:6689-92.

Cancer Research UK/NHS Centre for Statistics in Medicine, Wolfson College, Oxford OX2 6UD  
Douglas G Altman  
professor of statistics in medicine

Department of Health Sciences, University of York, York YO10 5DD  
J Martin Bland  
professor of health statistics

Correspondence to: Prof Altman  
doug.altman@cancer.org.uk

BMJ 2005;331:903