Answers

Statements *a* and *c* are true, whereas *b* is false.

The purpose of the study was to quantify the association between measurements of bone mineral density in twins, with monozygotic and dizygotic twins studied separately. Correlation is often used to quantify the association between pairs of measurements in a sample. Pearson’s or Spearman’s correlation coefficients, described in previous endgames, are used to quantify the linear association between two variables measured for each member of a sample. The coefficients might be used to measure the association between, for example, weight gain of low birthweight infants fed expressed breast milk and energy intake between the second and fourth weeks of life. Each pair of measurements is ordered—that is, one of the measurements in the pair relates to weight gain and the other measurement relates to energy intake.

In the example above it was not possible to use Pearson’s or Spearman’s correlation coefficients to quantify the association between twins in their lumbar spine bone mineral density measurements. Twins are unordered pairs and it would not be possible to uniquely assign each woman’s measurement to the pair of measurements required to derive Pearson’s or Spearman’s correlation coefficients. Changing the ordering of woman between sets of twins would greatly influence the value of these correlation coefficients. The pairs of measurements could be ordered in many ways. For example, there were 87 monozygoticsets of twins and each set’s measurements could be ordered in two ways in the pair; therefore the pairs of measurements for the sample could be arranged in 2^{87} (154<thin>742<thin>504<thin>910<thin>672<thin>534<thin>362<thin>390<thin>528) different ways.

The intraclass correlation coefficient was originally introduced in genetics to quantify the agreement between measurements for siblings (*a* is true). In effect, the intraclass correlation coefficients for lumbar spine bone mineral density were an average correlation coefficient across all possible orderings of measurements for twins in each sample. In particular, the intraclass correlation coefficient quantifies the agreement between pairs of measurements. This is in contrast to Pearson’s or Spearman’s correlation coefficients, which quantify the strength of the linear association between two measurements.

The derivation of the intraclass correlation coefficients for the samples of monozygotic and dizygotic twins was based on analysing the total variation in lumbar spine bone mineral density measurements for each sample. There are two sources for the total variation in measurements within a sample. Firstly, the variation between women within each set of twins aggregated across the sample; secondly, the variation in measurements between all sets of twins.

The intraclass correlation coefficient was derived as the ratio of the variation between all sets of twins to the total variation in lumbar spine bone mineral density measurements. Generally, as measurements within sets of twins show greater agreement, the variation between women within sets of twins aggregated across the sample will decrease. Therefore, as measurements between twins show greater agreement, the numerator and denominator of the intraclass correlation coefficient will become similar in value and the intraclass correlation coefficient will approach 1. Conversely, if the agreement between twins’ measurements were to decrease, the denominator would increase in value and the intraclass correlation coefficient would approach 0 in value. Unlike Pearson’s and Spearman’s correlation coefficients, which take a value from −1 through 0 to +1, the intraclass correlation coefficient takes a value from 0 to 1 (*b* is false). An intraclass correlation coefficient of 0 would demonstrate no agreement between twins in measurements of bone mineral density, whereas a coefficient of 1 would demonstrate perfect agreement. The coefficient is measured on a scale with no units.

The intraclass correlation of bone mineral density for the lumbar spine was higher for monozygotic than for dizygotic twins (0.74 *v* 0.36). Therefore, greater agreement existed between measurements of lumbar spine bone mineral density for monzygotic twins than for dizygotic ones (*c* is true). The authors of the article discussed that, because monozygotic twins are in complete agreement for genetic factors, any lack of agreement between twins would be due to environmental factors. Dizygotic twins share on average half of their genes. Therefore, any lack of agreement between dizygotic twins would be due to differences in both environmental and genetic factors. By comparing monozygotic and dizygotic twins, it was possible to estimate the extent to which genetic factors influence radiographic measures of osteoporosis.

As mentioned, the intraclass correlation coefficient was originally introduced in genetics to quantify the agreement between measurements for siblings. However, the intraclass correlation coefficient does not depend on having pairs of siblings. It can be used to study the agreement between measurements for two or more siblings, and the numbers of siblings in each family do not necessarily have to be the same. The intraclass correlation coefficient has several other important applications. It can be used to study measurement error and to calculate sample size in cluster randomised controlled trials. Both of these applications will be described in future endgames. Cluster randomised controlled trials have been described in a previous endgame.

Competing interests: none declared.

1. Spector TD, Keen RW, Arden NK, Morrison NA, Major PJ, Nguyen TV, et al. Influence of vitamin D receptor genotype on bone mineral density in postmenopausal women: a twin study in Britain. *BMJ* 1995;310:1357.
2. Sedgwick P. Pearson’s correlation coefficient. *BMJ* 2012;345:e4483.
3. Sedgwick P. Correlation. *BMJ* 2012;345:e5407.
4. Sedgwick P. Cluster randomised controlled trials. *BMJ* 2012;345:e4654.

Cite this as: *BMJ* 2013;346:f1816