thebmj

ENDGAMES

STATISTICAL QUESTION

Standardising outcome measures using z scores

Philip Sedgwick *reader in medical statistics and medical education*

Institute for Medical and Biomedical Education, St George's, University of London, London, UK

Researchers investigated if a shared care obesity management programme reduced body mass index (BMI) and related outcomes in obese children. A randomised controlled study design was used. Intervention consisted of general practice surveillance for childhood obesity, followed by obesity management across primary and tertiary care settings using a shared care model. Intervention was delivered over one year. Control consisted of "usual care." Participants were children aged 3-10 years with a BMI above the 95th centile for their age and sex. A total of 118 children were recruited through their general practice and randomised to intervention (n=62) or control (n=56).[1]

The main outcome was BMI. Measurements of BMI were transformed to z scores. Secondary outcome measures included body fat percentage, waist circumference, physical activity, quality of diet, and health related quality of life. At the end of follow-up, there was no significant difference between treatment groups in BMI (adjusted mean difference −0.1 (95% confidence interval −0.7 to 0.5; P=0.7)) and BMI z score (−0.05 (−0.14 to 0.03); P=0.2). No evidence of a significant difference was found for any secondary outcome. It was concluded the shared care model of primary and tertiary care management had no effect on BMI and related outcomes in obese children.

Which of the following statements, if any, are true?

a) The z scores permitted comparisons of BMI between children of different ages and sexes.

b) The z scores had the same units as BMI.

c) z scores are always positive in value.

d) The greater the z score, the heavier a child compared with other children of the same age and sex.

Answers

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

The aim of the study was to investigate if a shared care obesity management programme influenced BMI and related outcomes in obese children. A randomised controlled study design was used. The primary outcome was BMI. Measurements of BMI were transformed to z scores, sometimes called z values or standard deviation scores. The BMI z scores standardised the BMI measurements, allowing comparisons of BMI between children of different ages and sexes (*a* is true).

BMI is derived as weight in kilograms divided by the square of height in metres, and therefore is measured in the units of $kg.m^{-2}$. The index adjusts weight for height and provides a way of transforming weights at different heights into a common measure. BMI is an indicator of body fat, and is used to assess how much someone's body weight departs from what is normal or desirable for a person of his or her height. Fixed thresholds of BMI are used to describe obesity, healthy weight, and underweight in adults, regardless of their age. However, for children such thresholds vary with age and sex. The distribution of BMI is more complex for children than adults because it changes as children mature. Patterns of growth also differ between boys and girls. Therefore, in the trial above it would have been difficult to estimate the effect of treatment on BMI because the measurement would have depended on the age and sex of the child. The transformation of BMI into z scores standardised the BMI measurements, accounting for differences between children in age and sex and thereby allowing the effects of intervention on BMI to be evaluated.

In the study above, a child's BMI was compared with typical values for other children of the same age and sex. Reference data for the population of boys and girls aged between 3 and 10 years were provided by a large group of children who had previously had their weight and height measured. Children were recruited to the study above if their BMI was above the 95th centile of the distribution of BMI for their age and sex, and therefore by definition obese. Measurements of BMI for children in the study above were subsequently transformed into z scores using the reference data. For children of each age and sex in the reference data set, the distribution of BMI would have been described by a unique mean and standard deviation. The z score for a child in the study above was calculated as the difference between their BMI measurement and the population mean for their age and sex in the reference data set, divided by the standard deviation of the distribution of BMI for their age and sex. Therefore, a z score has no units because the numerator and denominator of the ratio are measured in the same units (*b* is false).

p.sedgwick@sgul.ac.uk

A z score describes how far away a child's BMI is from the population mean, expressed as a multiple of the population standard deviation. The value of a z score can be negative or positive depending on whether the child's BMI is smaller or greater than the population mean BMI for their age and sex (*c* is false). The further a child's BMI is away from the population mean for the child's age and sex, the larger the absolute value of his or her z score (ignoring the sign of the z score). Children were recruited to the study above if their BMI was above the 95th centile of the distribution of BMI for their age and sex. Therefore, their BMI would have been greater than the population mean BMI for their age and sex, and hence their z scores would have been positive in value. The greater the z score, the heavier a child compared with other children of the same age and sex (*d* is true).

Trials that investigate the effects of intervention on body mass and growth in children often have outcomes including several anthropometric measures made on different scales. In such studies the derivation of z scores not only permit a comparison between children of different ages and sexes for each outcome, but also a comparison of the effect of intervention between the anthropometric measures made on different scales. An example is given in a previous question.[2]

For the trial above, the treatment groups were compared in mean BMI z scores at the end of follow-up adjusted for other factors including baseline BMI z score. Analyses also included a comparison between treatment groups in mean BMI at end of follow-up, adjusted for children's age and sex at randomisation plus their baseline BMI. A multiple regression model, described in a previous question,[3] was used. Although BMI z scores were adjusted for age and sex, it was of benefit to compare the actual measurements of BMI and adjust for age and sex in a separate analysis. By undertaking analyses using different methods of measuring body mass it was possible to verify the observed effects of intervention on body mass index.

Competing interests: None declared.

1 Wake M, Lycett K, Clifford SA, Sabin MA, Gunn J, Gibbons K, et al. Shared care obesity management in 3-10 year old children: 12 month outcomes of HopSCOTCH randomised trial. *BMJ* 2013;346:f3092.
2 Sedgwick P. Standardisation of outcome measures (z scores). *BMJ* 2012;345:e6178.
3 Sedgwick P. Multiple regression. *BMJ* 2013;347:f4373.

Cite this as: *BMJ* 2014;22:g5878