BMJ 2015;351:h5119 doi: 10.1136/bmj.h5119 (Published 25 September 2015)

ENDGAMES

STATISTICAL QUESTION

Randomised controlled trials: understanding confounding

Philip Sedgwick *reader in medical statistics and medical education*

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

Researchers assessed the effectiveness of an exercise programme in reducing injurious falls among women at increased risk of falls and injuries. A multicentre parallel group randomised controlled trial study design was used. The intervention consisted of weekly supervised group sessions of progressive balance training offered in community based premises for two years, supplemented by individually prescribed home exercises. The control treatment was standard care. The study took place in 20 study centres throughout France. Participants were aged 75-85 years, living in their own home, and with reduced balance and gait capacities. In total, 706 women were recruited and randomly allocated to the intervention group (exercise programme; n=352) or control group (standard care; n=354). The random allocation of participants was stratified by study centre and body weight (<59 kg *v* ≥59 kg).[1]

The primary outcome was the rate of injurious falls (moderate and severe). Secondary outcomes included physical tests and perception of overall physical function. Of those women allocated to the intervention, 306 completed the trial protocol compared with 294 of those women allocated to control treatment. Analysis was by intention to treat. The risk of injurious falls was significantly reduced in the intervention group when compared with the control group (hazard ratio 0.81, 95% confidence interval 0.67 to 0.99). At two years, women in the intervention group performed significantly better on all physical tests and had significantly better perception of their overall physical function than women in the control group. It was concluded that the exercise programme was effective in reducing the risk of injurious falls, and in improving measured and perceived physical function in women aged 75-85 years at risk of falling.

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

a) The random allocation of the women to treatment group minimised confounding

b) Stratifying the random allocation of women controlled for the effects of study centre and body weight as potential confounders

c) The trial was prone to attrition bias

d) Intention to treat analysis minimised the effects of confounding

Answers

Statements *a*, *b*, *c*, and *d* are all true.

The aim of the trial was to assess the effectiveness of the exercise programme in reducing injurious falls among women at increased risk of falls and injuries. The participants were randomly allocated to treatment group. Therefore, a woman's characteristics did not influence which treatment she was allocated to and each woman had an equal probability of being allocated to each treatment group. Providing the sample size is large enough, random allocation results in treatment groups that are similar in baseline characteristics, thereby minimising confounding (*a* is true). Confounding is a difference between treatment groups in the distribution of those characteristics that influence the association between treatment and the outcome measures. These include demographic characteristics, prognostic factors, and other characteristics that may influence someone to participate in or withdraw from a trial. Therefore, if confounding is minimised at baseline, differences between the treatment groups in outcomes at the end of the trial will be due to differences in treatment and not to differences in baseline characteristics, thereby permitting the inference of causality to be ascribed to a treatment. If a characteristic is unequally distributed between the treatment groups at baseline and is associated with the treatment and outcome measure(s), it is referred to as a confounder.

The researchers considered study centre and body weight to be important prognostic factors that would influence the association between the treatment and the outcomes. In particular, the delivery of the treatment might be expected to differ between the study centres because different healthcare professionals were involved. Furthermore, body weight is a major risk factor for low bone mineral density and fractures in older women. However, simply randomly allocating women to treatment would not have guaranteed that the distribution of these prognostic factors would be similar in the treatment groups at baseline.

Correspondence to: P Sedgwick p.sedgwick@sgul.ac.uk

Hence, there was the potential for confounding by study centre and body weight. To minimise potential confounding by these important prognostic factors, the random allocation of women was stratified by study centre and body weight (<59 kg *v* ≥59 kg). Stratified random allocation has been described in a previous question.[2] Forty possible strata existed—for each of the 20 study centres there were two weight strata. Stratified random allocation involved allocating the trial participants within each stratum to treatment groups using simple random allocation. The aim was to achieve a similar distribution of the treatment groups in each stratum, thereby controlling for the potential confounding of the prognostic factors of study centre and body weight (*b* is true).

The trial was prone to attrition bias (*c* is true). Of the 352 women allocated to the intervention group, 306 (86.9%) completed the trial compared with 294 (83.1%) of the 354 women allocated to the control group. It is common for participants to drop out or be lost to follow-up in a trial. Attrition bias would have occurred if there had been a systematic difference between treatment groups in the characteristics of those participants who did not complete the trial or in their reasons for leaving the trial. Although attrition bias can be reduced or minimised by diligent follow-up, it is rarely eliminated. However, attrition bias is important only if the characteristics of those participants who did not complete the trial, or their reasons for leaving the trial, are associated with the outcome measures. Because not all of the women completed the trial and the women who dropped out did not provide outcome measures at two years, there was potential for confounding in the comparison of the treatment groups. Hence, the trial may not have had internal validity. Described in a previous question, internal validity is the extent to which the observed treatment effects can be ascribed to differences in treatment and not confounding, thereby allowing the inference of causality to be ascribed to the differences in treatment.[3]

The treatment groups were compared in the outcome measures at two years using an intention to treat analysis. Described in a previous question,[4] this approach compares treatment groups as originally allocated, irrespective of whether patients received or adhered to their treatment protocol. As described above, not all of the participants in the trial completed the trial. The intention to treat analysis promoted internal validity—it ensured that when the treatment groups were compared in outcome at follow-up they remained similar in baseline characteristics and it therefore minimised confounding (*d* is true).

If random allocation is successful and confounding is minimised, then treatment groups may be compared in the outcome measures using univariate statistical tests, such as the Student's *t* test and paired *t* test.[5] However, if confounding is present then it would be inappropriate to use such methods. It is possible to adjust for potential confounders during statistical analysis when examining the association between treatment and outcome measures. Such analytical approaches are referred to as multivariate and provide an estimate of the treatment effect, having adjusted for differences between treatment groups in the confounders. However, it is more efficient to adjust for confounding at the design stage of a study than to do so in subsequent analyses. Moreover, it may not be possible to adjust for all confounders because it is not always possible to measure all of them. Confounding is of particular concern when investigating the association between risk factors and outcome measures in observational studies—for example, cohort and case-control studies. Confounding in observational studies will be discussed in a later question.

Competing interests: None declared.

1 El-Khoury F, Cassou B, Latouche A, et al. Effectiveness of two year balance training programme on prevention of fall induced injuries in at risk women aged 75-85 living in community: Ossébo randomised controlled trial. *BMJ* 2015;351:h3830.
2 Sedgwick P. Treatment allocation in trials: stratified randomisation. *BMJ* 2015;350:h978.
3 Sedgwick P. Randomised controlled trials: internal versus external validity. *BMJ* 2014;348:g1742.
4 Sedgwick P. Intention to treat analysis versus per protocol analysis of trial data. *BMJ* 2015;350:h681.
5 Sedgwick P. Parametric statistical tests for independent groups: numerical data. *BMJ* 2012;345:e8145.

Cite this as: *BMJ* 2015;351:h5119.