BMJ 2015;351:h3869 doi: 10.1136/bmj.h3869 (Published 17 July 2015)

ENDGAMES

STATISTICAL QUESTION

Bias in experimental study designs: randomised controlled trials with parallel groups

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 the effects of antenatal lifestyle advice on health outcomes in overweight and obese pregnant women. A randomised controlled trial with a parallel groups study design was used. The intervention consisted of a comprehensive dietary and lifestyle programme during pregnancy, which incorporated a combination of dietary, exercise, and behavioural strategies delivered by a research dietitian and trained research assistants. The control treatment was standard pregnancy care, which did not include such strategies. Participants were women with a singleton pregnancy, between 10 and 20 weeks' gestation, and who had a body mass index (BMI) of 25 or more. The setting was three public maternity hospitals across South Australia.[1]

In total, 5474 eligible women were approached, of whom 2212 accepted the invitation and were randomly allocated to the intervention (n=1108) or control (n=1104) treatment. To ensure allocation concealment, a computer generated allocation sequence was delivered by a telephone randomisation service not involved with recruitment or clinical care. The primary outcome was the proportion of infants born large for gestational age (birth weight ≥90th centile for gestation and sex). Secondary outcomes included birth weight greater than 4000 g, hypertension, pre-eclampsia, and gestational diabetes. All outcomes were assessed by research staff blinded to the original treatment allocation.

Although the risk of infants being born large for gestational age was lower in the intervention group than in the control group, the difference was not significant (18.9% *v* 21%; adjusted relative risk 0.90, 95% confidence interval 0.77 to 1.07; P=0.24). Infants born to women after lifestyle advice were significantly less likely to have a birth weight above 4000 g (15.3% *v* 18.8%; 0.82, 0.68 to 0.99; P=0.04). There were no differences between the treatment groups for any of the other secondary outcomes. It was concluded that, in overweight or obese women, antenatal lifestyle advice did not reduce the risk of a baby being born large for gestational age or improve maternal pregnancy and birth outcomes.

Which of the following, if any, might the above trial have been prone to?

a) Allocation bias

b) Ascertainment bias

c) Non-response bias

d) Resentful demoralisation

e) Selection bias

f) Volunteer bias

Answers

Answers *c*, *d*, *e*, and *f* are true, whereas *a* and *b* are false.

The aim of the trial was to investigate the effectiveness of antenatal lifestyle advice on health outcomes in overweight and obese pregnant women. A randomised controlled trial with a parallel groups study design was performed. Participants were allocated to the intervention or control treatment, which they received for the entire study period. The treatment groups were followed alongside each other—hence the term parallel groups study design. The design is sometimes referred to as a "between subjects" design because the outcomes are compared between independent groups of patients—that is, between subjects.

In the above trial, the participants were women with a singleton pregnancy, between 10 and 20 weeks' gestation, who had a BMI of 25 or more. The women were recruited from three public maternity hospitals across South Australia. As for any study, the method of sampling and availability of potential participants will have affected the extent of selection bias. Selection bias would have occurred if the sample was systematically different from the population it was intended to represent. In this trial, the sample was obtained using convenience sampling,[2] with women recruited from three maternity hospitals in South Australia. The sample would therefore probably have been systematically different from the population with respect to its demographics and health outcomes. Hence the sample would be prone to selection bias (*e* is true). If selection bias exists it results in a lack of external validity—that is, the extent to which the study results can be generalised to the population that the sample is meant to represent.[3]

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

Selection bias is a general term used to describe a group of biases and effects that result in a sample that is systematically different from the population it is intended to represent. Non-response bias and volunteer bias belong to this group and are described below.

In total, 5474 women were eligible and invited to take part in the trial, and 2212 (40.4%) accepted the invitation. Therefore, the sample was prone to non-response bias (*c* is true). Non-response bias would have occurred if the women who accepted the invitation were systematically different from those who did not. The women who accepted the invitation would have been expected to be different from those who did not, not least in their motivation to take part in the trial. However, any differences in characteristics (including sociodemographics and prognostic factors) would be difficult to quantify because limited information, if any, would be available for those who refused to be part of the trial. Non-response bias should not be confused with response bias,[4] or volunteer bias.

The sample in the above trial was prone to volunteer bias (*f* is true)—a systematic difference between those people who volunteered to be part of the trial and the population. The volunteers would be expected to differ from the population in their sociodemography, behaviour, attitudes, and health. It has been reported that, in general, those who participate in studies are more educated, come from a higher social class, and are more sociable than those who do not volunteer. Non-response bias and volunteer bias are often confused. Non-response bias focuses on the potential differences between the non-responders and responders originally invited to be part of the sample, whereas volunteer bias considers the potential differences between those who volunteered and the population. Both non-response bias and volunteer bias result in selection bias.

Allocation bias would have occurred if there was a systematic difference between participants in how they were allocated to treatment groups. For example, researchers may have favoured the intervention, wishing to show that it was more effective than the control treatment, and allocated those women to the intervention whom they believed would have experienced the greatest benefit from the treatment. In the above trial it was not possible to blind participants to their treatment allocation. However, allocation concealment[5] was achieved by the use of an independent telephone randomisation service not involved with recruitment, clinical care, or assessment of the outcomes. Therefore, the allocation sequence was not disclosed to women and those recruiting the trial participants. The order in which participants were to be allocated to treatments was random, so participants had an equal probability of being allocated to each treatment group.[6] Hence, allocation bias was eliminated (*a* is false).

The use of random allocation and the elimination of allocation bias meant that treatment groups would be similar in baseline characteristics, thereby minimising confounding. Confounding is a difference between treatment groups in the characteristics that influence the association between the treatments and outcomes. These include demographic characteristics, prognostic factors, and other characteristics that may influence someone to participate in or withdraw from a trial. Therefore, if confounding was minimised at baseline, any differences between the treatment groups in outcomes at the end of the trial would be due to differences in treatment and not to differences in baseline characteristics.

Ascertainment bias, sometimes referred to as detection bias, would have occurred if the recorded measurements of the pregnancy and birth outcomes were systematically different from those experienced by the women. Such bias in data collection can be unconscious or otherwise and can originate from the investigators or participants. When ascertainment bias occurs on behalf of the participants, it is referred to as response bias. If ascertainment bias occurs on behalf of the researchers it is referred to as assessment bias, sometimes known as observer bias. Blinding of the participants and outcome assessors is necessary to minimise ascertainment bias. However, in the above trial it was not possible to blind participants to their treatment allocation. Nonetheless, the outcome measures were not self reported by the participants, so response bias would not have occurred. If the investigators were aware of the treatment allocation, they could have been biased in their assessment of a treatment group. Assessment bias would have occurred, for example, if the researchers favoured the intervention and wished to show that it was more effective than the control treatment. However, because all outcomes were assessed by research staff blinded to the original treatment allocation, it is unlikely that the outcome measurements were prone to ascertainment bias (*b* is false) or, more specifically, assessment bias.

It was not possible to blind participants to their treatment allocation in the above trial. It is therefore possible that the women's responses to treatment might have been influenced by knowledge of their treatment allocation. Some participants might have had a preference for one of the treatments—the intervention or control. Women who received their preferred treatment were more likely to have been better motivated and shown greater adherence to their treatment regimen. By contrast, women who did not receive their preferred treatment might have exhibited resentful demoralisation (*d* is true), whereby they complied poorly and possibly withdrew from the trial.[7] If women had a preference for a treatment it might also have contributed to performance bias, where factors other than the treatments under investigation influenced the outcome measures. In particular, a woman's preference for a particular treatment might have had an important effect on the perceived benefit of and reporting of side effects for the treatment she was allocated.

The above trial was a between subjects study design, whereby each women received just one treatment throughout the entire study. Some trials have a "within subjects" study design, whereby each participant receives all available treatments, the order of which is determined at random.[8] Such designs are prone to other biases in addition to those described above, and these will be described in a future question.

Competing interest: None declared.

1 Dodd JM, Turnbull D, McPhee AJ, et al; for the LIMIT Randomised Trial Group. Antenatal lifestyle advice for women who are overweight or obese: LIMIT randomised trial. *BMJ* 2014;348:g1285.
2 Sedgwick P. Convenience sampling. *BMJ* 2013;347:f6304.
3 Sedgwick P. Randomised controlled trials: internal versus external validity. *BMJ* 2014;348:g1742.
4 Sedgwick P. Non-response bias versus response bias. *BMJ* 2014;348:g2573.
5 Sedgwick P. Controlled trials: allocation concealment, random allocation, and blinding. *BMJ* 2015;350:h2633.
6 Sedgwick P. Why randomise in clinical trials? *BMJ* 2012;345:e5584.
7 Sedgwick P. What is a patient preference trial? *BMJ* 2013;347:f5970.
8 Sedgwick P. Randomised controlled trials: "within subject" versus "between subject" designs. *BMJ* 2014;349:g6435.

Cite this as: *BMJ* 2015;351:h3869