BMJ

BMJ 2011;343:d7453 doi: 10.1136/bmj.d7453 (Published 23 November 2011)

ENDGAMES

STATISTICAL QUESTION

Random sampling versus random allocation

Philip Sedgwick *senior lecturer in medical statistics*

Centre for Medical and Healthcare Education, St George's, University of London, Tooting, London, UK

Researchers investigated the effects of providing people with evidence based information about colorectal cancer and screening. A randomised controlled study design was used. The intervention was a brochure that included personalised risk of colorectal cancer, available screening options with possible benefits and harm, plus information on prevention of colorectal cancer. People also had access to two optional interactive internet modules on risk and diagnostic tests. The control treatment was the official information leaflet of the German colorectal cancer screening programme. Primary outcome measures included "informed choice" based on knowledge and attitude, plus the planned and actual uptake of screening.[1]

Potential trial participants were people insured by a large German statutory health insurance scheme who were aged 50-75 years and had no history of colorectal cancer. A total of 7946 people were eligible. A random sample of about 4000 people was drawn and invited to participate in the trial. In total, 1577 agreed and were randomly allocated to the intervention (n=785) or control (n=792).

The researchers reported that providing evidence based risk information on colorectal cancer and screening improved knowledge and increased informed choices but had little effect on attitudes. The intervention did not increase combined actual and planned uptake of screening.

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

a) The objective of random sampling was to facilitate application of the trial results to the study population

b) Random allocation of trial participants minimised allocation bias

c) Random sampling ensured treatment groups were of similar size

d) Random allocation of trial participants minimised confounding at baseline

e) Random allocation of trial participants facilitated application of the trial results to the study population

Answers

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

The methods of random sampling and random allocation are often confused. Random sampling was the method by which participants were selected for the trial, whereas random allocation was how participants were allocated to treatment groups.

The study population is the entire group of people of interest, although it is not always well defined. For the above trial, the study population comprised people living in Germany who were aged 50-75 years and had no history of colorectal cancer. It was assumed that people who belonged to the insurance scheme were representative of the German population. Simple random sampling, often referred to as random sampling, was used to obtain a sample of the study population from the insurance scheme. The objective was to obtain a representative sample of people in the insurance scheme in all characteristics, including demography and disease severity, so that the study results could be applied to the study population (*a* is true). A sampling frame was constructed—that is, a list of all people belonging to the insurance scheme aged 50-75 years who had no history of colorectal cancer. A sample of fixed size was selected at random from this list, with all people having the same probability of being selected independently of all others. Random samples will be representative of the study population if they are large enough.

Trial participants were allocated to the intervention or control using simple random allocation, often referred to as random allocation or randomisation. Each person was allocated at random and, therefore, had an equal probability of being allocated to the intervention or control. The characteristics of the trial participants did not influence which treatment group they were allocated to, so allocation bias was minimised (*b* is true). Allocation bias would have occurred if the characteristics of those people allocated to the intervention differed from those allocated to the control. Participants were allocated in a 1:1 ratio—on average, for every one person allocated to the intervention one would be allocated to the control. Random allocation therefore achieved treatment groups of similar size. Random sampling is not how participants were allocated to treatment groups, but the method of selecting participants for the trial (*c* is false).

p.sedgwick@sgul.ac.uk

Random allocation aimed to minimise confounding by achieving comparability of treatment groups in baseline characteristics (*d* is true). Confounding is the difference between treatment groups in baseline characteristics that influence treatment and outcome measures. These factors include demographics, prognostic factors, and other characteristics that may influence someone to participate in or withdraw from a trial. If confounding is minimised, any differences between treatment groups in outcome at the end of the trial will result from differences in treatment and not differences in baseline characteristics. Random allocation achieves comparability in baseline characteristics only if the sample size is large enough.

The objective of random sampling was to obtain a representative sample of the study population. By randomly allocating the sample to the intervention or control, the two treatment groups would be similar in baseline characteristics and similar to the original sample. Random allocation therefore ensured that the study results could be applied to the study population (*e* is true). In this respect, random allocation and random sampling have the same objectives.

Competing interests: None declared.

1 Steckelberg A, Hülfenhaus C, Haastert B, Mühlhauser I. Effect of evidence based risk information on “informed choice” in colorectal cancer screening: randomised controlled trial. *BMJ* 2011;342:d3193.

Cite this as: *BMJ* 2011;343:d7453