

ENDGAMES

STATISTICAL QUESTION

External and internal validity in clinical trials

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 efficacy of a probiotic drink containing *Lactobacillus* for the prevention of antibiotic associated diarrhoea in patients over 50 years. A randomised double blind placebo controlled trial study design was used. The intervention consisted of consumption of a probiotic drink twice a day during a course of antibiotics and for one week after the course finished. The placebo was a long life sterile milkshake. The primary outcome was occurrence of antibiotic associated diarrhoea.¹

Participants were recruited from three London hospitals. A total of 1760 patients were assessed for eligibility, 1625 of whom were not recruited because they did not meet the inclusion criteria (n=1263), refused to participate (n=148), or could not be included for practical reasons (n=214). The remaining 135 patients were recruited to the trial and randomised to intervention (n=69) or placebo (n=66). In total, 12 patients receiving the intervention and 10 in the placebo group did not complete their treatment protocol because they were lost to follow-up, withdrew consent, or died during the study period. The trial was analysed using a per protocol analysis. The researchers reported that consumption of the probiotic drink reduced the incidence of antibiotic associated diarrhoea.

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

- The random allocation of patients promoted external validity
- The random allocation of patients promoted internal validity
- The per protocol analysis promoted internal validity

Answers

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

External validity and internal validity are essential components in the design of clinical trials. External validity is the extent to which the study results can be generalised to a population, in particular the population the sample is meant to represent. Internal validity is the extent to which observed treatment effects can be ascribed to differences in treatment and not confounding, thereby allowing the inference of causality to be ascribed to a treatment.

The above population consisted of hospital patients aged over 50 years although it was limited by a series of exclusions. Patients were excluded if they had diarrhoea on admission, bowel pathology that could result in diarrhoea, severe illness, immunosuppression, bowel surgery, artificial heart valves, a history of rheumatic heart disease or infective endocarditis, or if they had used antibiotics in the previous four weeks. Of the original 1760 patients assessed, 497 were eligible to take part in the trial. It is assumed that these 497 participants were representative of the population. However, only 113 of the potential participants were eventually recruited. It is not clear whether the final sample was representative of the population, so it is difficult to assess the extent of external validity. The patients recruited to the trial may differ from those who were eligible but not recruited in terms of sociodemographic and prognostic factors. As the proportion of eligible patients who refuse to participate in a trial increases, external validity decreases.

The purpose of random allocation is to promote external validity and internal validity (*a* and *b* are true). The patients in the sample were randomly allocated to treatment or placebo to achieve two groups that were similar in baseline characteristics. In turn, the treatment groups should have similar characteristics to the sample taken from the population, thereby promoting external validity (answer *a*). Confounding—differences in baseline characteristics between treatment groups that influence treatment and outcome measures—is minimised if the two treatment groups are similar in baseline characteristics. When confounding is minimised, any differences between treatment groups in outcome at the end of the trial will result from differences in treatment and not from differences in baseline characteristics (answer *b*). Random allocation achieves greater comparability in baseline characteristics as sample size increases. This trial used a restricted randomisation process, whereby participants were allocated to treatment using an allocation sequence stratified for hospital, sex, and age group (50-69 and over 70). By doing so, the random allocation procedure achieved greater equivalence between treatment groups in group sizes and baseline characteristics.

Per protocol analysis, described in a previous question,² compares treatment groups as originally allocated but includes only those patients who completed the treatment protocol. In this study, 12 patients in the intervention group and 10 in the placebo were lost to follow-up, withdrew consent, or died. These patients would not have been included in the analysis because they did not complete the treatment protocol. Therefore, the analysis of the treatment groups did not promote internal validity (*c* is false) because the similarity in baseline characteristics achieved after randomisation may have been compromised. The rationale behind using a per protocol analysis is that the estimated efficacy of the intervention is not affected by factors such as non-compliance. Because of the potential for biased results, patients who were excluded should be carefully considered and described. In contrast, an intention to treat

analysis, described in a previous question,³ compares treatment groups as originally allocated, irrespective of whether patients received or adhered to their treatment protocol. Intention to treat analysis promotes external validity because it is a pragmatic approach that aims to evaluate the effectiveness of an intervention in routine practice.

Competing interests: None declared.

- 1 Hickson M, D'Souza AL, Muthu N, Rogers TR, Want S, Rajkumar C, et al. Use of probiotic Lactobacillus preparation to prevent diarrhoea associated with antibiotics: randomised double blind placebo controlled trial. *BMJ* 2007;335:80.
- 2 Sedgwick P. Analysis by per protocol. *BMJ* 2011;342:d2330.
- 3 Sedgwick P. Analyses by intention to treat. *BMJ* 2011;342:d2212.

Cite this as: *BMJ* 2012;344:e1004

© BMJ Publishing Group Ltd 2012