

ENDGAMES

STATISTICAL QUESTION

Why match in case-control studies?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 effectiveness of a monovalent rotavirus vaccine against severe rotavirus diarrhoea in children aged under 2 years. A case-control study design was used. In total, 323 children were recruited from seven hospitals in El Salvador after admission with laboratory confirmed rotavirus diarrhoea. For each case three controls, matched for age (within 30 days of date of birth) and neighbourhood, were recruited.¹

Vaccination history was confirmed after inspection of vaccination cards held by the parents. Potential risk factors—including demographics, socioeconomic factors, birth weight, premature birth, current body mass index, history of breast feeding, day care attendance and medical history—were collected from hospital records or in interviews with parents. The researchers concluded that the monovalent rotavirus vaccine was highly effective against admissions for rotavirus diarrhoea in children aged under 2 years in El Salvador. No differences were reported between cases and controls in breastfeeding patterns, premature birth, maternal education, or socioeconomic variables.

Which of the following types of bias, if any, would matching of cases and controls have minimised?

- Allocation bias
- Confounding
- Ecological fallacy
- Recall bias

Answers

Confounding (answer *b*) was minimised by matching of cases and controls, whereas allocation bias, the ecological fallacy, and recall bias (answers *a*, *c*, and *d*) were not.

This case-control study aimed to evaluate the effectiveness of a monovalent rotavirus vaccination in reducing admissions for rotavirus diarrhoea in children aged under 2 years in El Salvador while also identifying potential risk factors. Two groups of children were selected on the basis of their disease status: the cases—those admitted to hospital with laboratory confirmed rotavirus diarrhoea—and healthy control children. Therefore, allocation bias would not have occurred (answer *a*). Allocation bias is of concern in clinical trials, it being the systematic

difference between how participants are allocated to treatment. Allocation bias has been described in a previous question.²

The controls were matched to cases with regard to age and neighbourhood in a ratio of 3:1. For each case recruited, three controls of the same age (born within 30 days of the case) who lived in the same neighbourhood had to be found. The researchers visited homes to the left and right of the case's home until three controls were identified, and recruitment was limited to one control per household. The purpose of matching was to minimise confounding by reducing systematic differences between the groups of cases and controls (answer *b*). Any differences between cases and controls would therefore not result from differences in age or neighbourhood, but rather differences in rotavirus vaccination status and other potential risk factors.

Because cases and controls were matched for age and neighbourhood, the relation between these two variables and rotavirus diarrhoea could not be examined. Although age and neighbourhood may be related to the development of rotavirus diarrhoea, they were not of interest as potential risk factors in this study. Matching cases and controls for age and neighbourhood meant that these potential confounders were adjusted for when the study was designed, which is more efficient than making adjustments during statistical analyses.

The ecological fallacy, which was described in a previous question,³ could not have occurred in the case-control study (answer *c*). The ecological fallacy typically occurs when results from an ecological study are being interpreted. In an ecological study data are aggregated for groups of people—for example, towns, cities, or countries. Associations between risk factors and diseases are then investigated for the groups of people, allowing an initial examination of the health status and needs of communities. The ecological fallacy is a bias that occurs when conclusions from an ecological study are mistakenly assumed to apply to individual people within the population. Although children in the case-control study were recruited from one of seven hospitals in El Salvador, and controls were subsequently matched to cases in the neighbourhood where they lived, the data were collected and analysed for individual children and not groups of children.

Matching cases and controls will not have reduced recall bias (answer *d*). Information recall may be inaccurate when past behaviour or exposure to potential risk factors is being reported. Recall bias is the systematic difference between cases and controls in the accuracy of reported information. All parents were interviewed about history of breast feeding, day care attendance, birth weight, and socioeconomic status. However, parents of cases may have recalled past behaviour more accurately than parents of controls because of the worry of having a child admitted to hospital with severe rotavirus diarrhoea.

Competing interests: None declared.

- 1 De Palma O, Cruz L, Ramos H, de Baires A, Villatoro N, Pastor D, et al. Effectiveness of rotavirus vaccination against childhood diarrhoea in El Salvador: case-control study. *BMJ* 2010;341:c2825.
- 2 Sedgwick P. Bias in clinical trials. *BMJ* 2011;342:d4176.
- 3 Sedgwick P. The ecological fallacy. *BMJ* 2011;343:d4670.

Cite this as: *BMJ* 2012;344:e691

© BMJ Publishing Group Ltd 2012