

Epidemiology 2

Compared to what? Finding controls for case-control studies

David A Grimes, Kenneth F Schulz

Lancet 2005; 365: 1429–33

Family Health International,
PO Box 13950, Research
Triangle Park, NC 27709, USA
(D A Grimes MD, K F Schulz PhD)

Correspondence to:
Dr David A Grimes
dgrimes@fhi.org

Use of control (comparison) groups is a powerful research tool. In case-control studies, controls estimate the frequency of an exposure in the population under study. Controls can be taken from known or unknown study populations. A known group consists of a defined population observed over a period, such as passengers on a cruise ship. When the study group is known, a sample of the population can be used as controls. If no population roster exists, then techniques such as random-digit dialling can be used. Sometimes, however, the study group is unknown, for example, motor-vehicle crash victims brought to an emergency department, who may come from far away. In this situation, hospital controls, neighbourhood controls, and friend, associate, or relative controls can be used. In general, one well-selected control group is better than two or more. When the number of cases is small, the ratio of controls to cases can be raised to improve the ability to find important differences. Although no ideal control group exists, readers need to think carefully about how representative the controls are. Poor choice of controls can lead to both wrong results and possible medical harm.

In an old movie, comedian Groucho Marx is asked: “Groucho, how’s your wife?” Groucho quips: “Compared to what?” Although sexist by contemporary standards, Groucho’s reply frames the question relating to the results of case-control studies: compared to what? Valid conclusions hinge on finding an appropriate comparison group. Stated alternatively, use of suboptimal control groups has undermined much research.

Use of control groups is a powerful scientific tool—and an old one. The first documentation of a comparison group appears in *The Holy Bible* in the Book of Daniel.¹ Daniel and his three colleagues, captured by King Nebuchadnezzar of Babylon, carried out a 10-day trial of healthy food versus the royal diet of the court. At the end, Daniel and his buddies appeared healthier than did the Babylonian youth who enjoyed the usual fare. This trial has been criticised over the years for both an inadequate duration of exposure and probable divine cointervention. Perhaps as a result, control groups disappeared from published work for millennia.

James Lind’s 1747 trial of scurvy treatments rekindled interest in contemporaneous controls.² Despite its small size (six treatment groups with two sailors assigned to each), the trial showed the benefit of citrus-fruit supplementation. In studies without randomisation, finding an appropriate control group can sometimes be challenging. We will explain the role of control groups in case-control studies, describe special difficulties in choosing them, and discuss some implications of these choices.

Aim of controls

Controls in a case-control study, which progresses backwards in time from outcome to exposure,³ indicate the background frequency of an exposure in individuals who are free of the disease in question. Controls do not need to be healthy; inclusion of sick people is sometimes appropriate. Indeed, exclusion of ill people as controls

can distort (bias)⁴ the results.⁵ (Like healthy individuals, ill people can develop a different condition of interest.) The final point is key: controls in a case-control study should represent those at risk of becoming a case.⁶ Stated another way, controls should have the same risk of exposure as the cases, if the exposure and disease are unrelated (panel).⁷

If cases (with the disease) have a higher frequency of the exposure than do the controls, then a positive association emerges—eg, multiple sexual partners are more common among cases of cervical cancer than among controls without cervical cancer. If the exposure prevalence among cases is lower than among controls, a protective association exists—eg, oral contraceptive use is less common among ovarian cancer cases than among controls without this cancer.

Avoidance of bias is important when choosing controls for a case-control study. Selection bias arises if controls are not representative of those at risk of the disease in question. Case-control studies of potential protection against colorectal cancer associated with non-steroidal anti-inflammatory drugs (NSAIDs)^{8,9} are a good example (figure 1). Assume that colorectal cancer cases are identified at the time of their operations in hospital. Controls are to be hospital patients without colorectal cancer. If we identified controls from the rheumatology service, this selection would bias the results, because patients with arthritis would be more likely than most people in the community to be exposed to NSAIDs, thereby reducing the estimate of the association between these drugs and colorectal cancer. By contrast, if controls

Panel: Attributes of controls in a case-control study

- Free of the outcome of interest
- Representative of the population at risk of the outcome
- Selected independent of the exposure of interest

Control selection	Flaw in control group	Effect on estimate of NSAID effect
Patients with arthritis	Probable atypically high exposure to NSAIDs	↓
Patients with peptic ulcers	Probable atypically low exposure to NSAIDs	↑

Figure 1: Introduction of bias in a case-control study of non-steroidal anti-inflammatory drugs (NSAIDs) and colorectal cancer
Cases are hospitalised patients with colorectal cancer.

were selected from the gastroenterology service, this choice would bias the results in the opposite direction. Patients with ulcers would be less likely than the general population at risk of colorectal cancer to be exposed to NSAIDs, because of warnings from their clinicians. This bias would increase the estimate of the effect.³

Research in endometriosis provides another example of challenges in selection of a control group. Since endometriosis needs an operation for diagnosis, investigators frequently use as controls women having laparoscopy or laparotomy without this diagnosis being made. However, women having operations are unlikely to be representative of all those at risk of developing endometriosis, since operations do not occur at random.¹⁰

Where to find controls?

The investigator (and, ultimately, the reader) needs to determine the group of individuals from which cases and controls will be drawn. A known group¹¹ consists of a defined population observed over a period (figure 2). This group might consist of passengers and crew on a week-long cruise of the Caribbean or all individuals living in Sweden over a decade. Cases are those who develop the disorder of interest and controls are those in the same group without the condition. Thus, case-control studies can be thought of as occurring in the midst of a larger cohort study (nested case-control studies¹² being a nice example). The task here is to find the cases in the group in question; choosing controls is easier in a defined population.

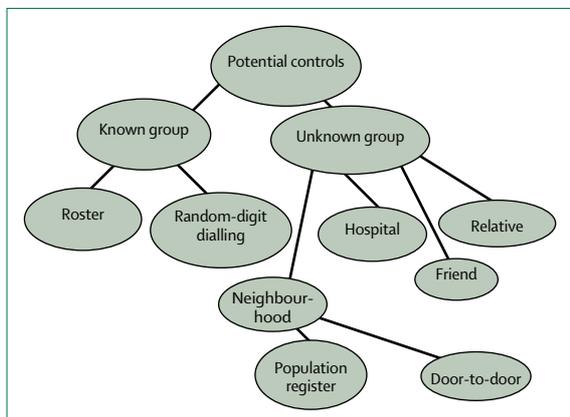


Figure 2: Choosing controls with known and unknown group of study participants

Usually the group from which cases come is unknown.¹¹ For example, victims of motor-vehicle crashes in a hospital emergency department pose this sort of challenge. Some might live nearby, others could be passing through on a highway, and others may arrive from rural areas by helicopter. Here, the cases are chosen before the study group is deduced. Finding cases is the simple part; the challenge now is to define the group from which controls should come. They should come from the same group. (One approach would be to limit cases and controls to people who live within the city limits.)

Poor control groups can lead to big mistakes. For example, an early case-control study of AIDS in homosexual men in San Francisco used two control groups: neighbourhood and clinic.¹³ The odds ratios for the exposure (>100 lifetime partners) were 52.0 for the neighbourhood controls and 2.9 for the clinic controls. Use of clinic controls grossly underestimated the true risk, since the likelihood of using a clinic was strongly related to the exposure of interest—ie, it was not independent of the number of partners. Controls in a public clinic at which sexually transmitted diseases are treated were much more likely to have multiple partners than were other homosexual men in San Francisco.

Controls from a known group

When possible, random samples of people without the disease can serve as controls. Investigation of an outbreak of food-borne illness on a cruise ship generally uses a case-control approach. Cases are those who develop gastroenteritis; controls are those on board who do not. The study seeks to identify food exposures that are more common among the cases than the controls. Moreover, no one who had not eaten the suspect food should have become ill. On the ship, probability sampling among those unaffected could be done.¹⁴ Thus, controls could be a random sample of everyone on board without gastroenteritis.

Population controls have both advantages and disadvantages. Random sampling should provide representative controls, and extrapolation of results to the study group is easily justified. On the other hand, population controls can be inappropriate when cases have not been completely identified in the population or when substantial numbers of potential controls cannot be reached—eg, those on holiday. Moreover, population controls could be less motivated to take part in research than individuals in a health-care setting, such as hospitalised patients.¹⁵

When no roster of the population exists, random-digit dialling of telephone numbers enables sampling of potential controls.^{7,16} A random sample of incomplete telephone numbers (eg, eight digits) is taken from working telephone exchanges; random two-digit numbers then complete the number to be called (figure 3). This approach has strengths and weaknesses.

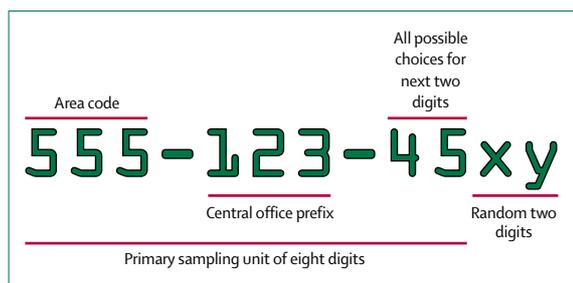


Figure 3: Random-digit dialling for controls¹⁷

Primary sampling unit included eight-digit numbers: all known area codes and three-digit central-office prefixes in the county, plus all combinations of next two digits. For all these eight-digit numbers randomly chosen, a computer generated the two final digits, creating a ten-digit number to be called.

It attempts to sample residential telephone numbers equally while keeping calls to commercial numbers to a minimum. The strategy reaches both new numbers and unlisted numbers not available through directories. Although it provides a random sample of telephone numbers, a random sample of potential controls is the real goal. Not all people have telephones; those without tend to be of lower socioeconomic status. Moreover, some individuals have more than one telephone number (eg, home plus cellular telephone), which could be related to higher socioeconomic status. Such people have an increased likelihood of being contacted. Some telephone numbers are used by more than one potential control. Individuals who are reluctant to respond to telephone enquiries differ from those who readily agree to participate.¹⁸ For example, young women are less likely to be found by random-digit dialling than are others.¹⁹ Although these quirks of telephone coverage might lead to bias,¹⁵ control groups selected this way are largely representative of the reference population.¹⁷

Controls from an unknown group

Neighbourhood controls

Neighbourhood controls generally are drawn in a specified pattern from the block in which the case lives. As always, selection of controls should be independent of the exposure of interest. To avoid selection bias, interviewers are given a specific pattern of houses to approach. Researchers have used two approaches to identify houses of controls: a population register or door-to-door canvassing.^{20,21} A useful aid for the former is the cross-reference (also termed crisscross or reverse-street) directory that lists addresses and corresponding telephone numbers.

A case-control study of oral contraceptives and hepatocellular adenomas used door-to-door canvassing; researchers interviewed every case in her home and then attempted to find three controls on the same street (figure 4).²² Another case-control study in a rural setting mapped the case's neighbourhood. Interviewers began by facing the case's house and then setting out in increasing circles of houses until an appropriate match was found.²¹

Advantages of neighbourhood controls include no need for a roster and that many confounding factors are accounted for—eg, socioeconomic status, climate, etc). On the other hand, canvassing neighbourhoods is expensive²¹ and using homes rather than people as the sampling unit is a problem shared with random-digit dialling. Non-response can pose challenges. In one report, an average of nine household contacts was needed for one successful control,²⁰ although in our experience this ratio can be as much as 150/1. Multiunit buildings require identification of all units and then gain of access. This challenge is not unique to urban settings; in a case-control study in which we participated, interviewers dealt with German Shepherd dogs, barbed wire fences, and arrest by suspicious local police.²²

Hospital controls

Hospital controls have been widely used—and criticised—in case-control studies. They have several appealing features: convenience, low-cost to identify and interview, comparable information quality as cases, motivation to participate, and comparable health-care-seeking behaviour.¹⁵ However, the disadvantages are notable. Use of hospital controls assumes that they are representative of the background rate of exposure among people in the study group that produced the cases, meaning that the exposure is unrelated to the disease leading to hospitalisation of the control. The best way to avoid this pitfall is to exclude as controls those whose admission diagnosis is likely to be related to the exposure of interest. For example, in a hospital-based case-control study of contraception and systemic lupus erythematosus, controls admitted to the obstetrics and

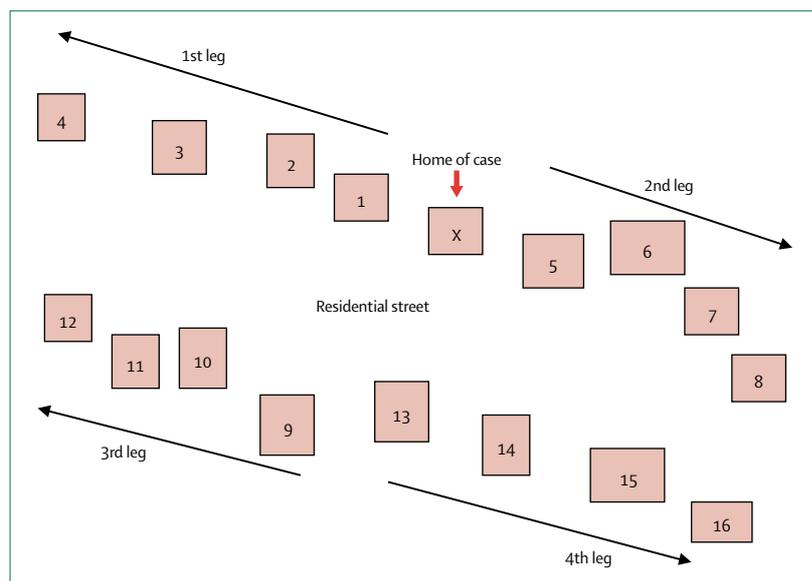


Figure 4: Neighbourhood controls

After interviewing the case in her home, the investigator canvasses up to 16 homes in a predetermined H-shaped pattern until three controls are identified and interviewed. Every rectangle represents a home along the same street.

gynaecology service were excluded.²³ The rationale for this exclusion was that women having reproductive care at tertiary-care hospitals might have had obstetric and gynaecological histories different than most women in the community. Different diseases can have different catchment areas for a hospital; control diagnoses should have the same catchment area as the cases.

Admission rate bias can also cause difficulties.^{4,24} For example, if women wearing an intrauterine device are more likely to be admitted for treatment of salpingitis than are women with salpingitis but no device, this difference would exaggerate the apparent odds ratio of salpingitis associated with intrauterine device use.²⁵

Several reports suggest that hospital controls might not be representative of the study group. Hospital controls can resemble cases more than do population controls,^{26,27} and others have noted substantial differences between hospital and population controls in weight, smoking patterns, and burden of illness (affecting the probability of hospitalisation).²⁸

Friend or associate controls

Friends or work associates of cases sometimes serve as controls. This approach has both critics and supporters. An advantage is generation of a control group similar to the cases in several important respects—such as socioeconomic status and education. However, asking cases to name potential controls is the antithesis of random selection. Those named might be more gregarious and sociable than other potential controls, leading to the controls not being representative.¹⁵ On the other hand, in hidden populations for which socially unacceptable behaviours are being studied, eg, drug abuse, friend controls have been suggested to be convenient and unlikely to introduce selection bias. In one study, drug misusers were asked to nominate a friend who was a drug misuser (a new case) and another friend who had never been involved with drugs (a control). This chain referral or snowball technique concluded that cases and controls came from the same population.²⁹

Relative controls

Relatives share many traits with cases. When genetic factors are deemed to be confounding, relatives have been used to control for this bias.¹⁵ Many other exposures will be similar—eg, siblings are likely to have diet, environment, lifestyle, and socioeconomic status in common as well. For example, when siblings serve as controls, the potential effect of family size cannot be examined.¹⁵ Some researchers have concluded that as long as the exposure-specific risks remain stable over time, use of relatives as controls does not distort the results.³⁰

How many control groups?

Some authors have argued for using two separate control groups; if results are consistent, then findings are deemed more credible.^{14,31} For example, a case-control study of

oestrogen therapy and endometrial cancer used both hospital and community controls.³² In the unhappy circumstance of disparate results, however, which result should be ignored?¹⁵ Another immediate disadvantage is the added cost in time and resources. For example, in the case-control study of endometrial cancer cited above,³² adding a community control group increased the number of study participants to be interviewed from 480 to 801, a 67% rise. In general, we suggest selection of the best control group possible.³³

How many controls per case?

Readers are sometimes surprised to discover large disparities between the numbers of cases and controls in a case-control report; clinicians intuitively expect similar group sizes.³⁴ This inequality reflects attempts by investigators to increase the ability of the study to find differences of importance, should they exist. In unmatched case-control studies, having roughly equal numbers of cases and controls is most efficient if costs are similar for cases and controls. However, sometimes the number of cases is small and cannot be increased—eg, 11 cases of liver cancer among young women in Los Angeles county over 5 years.³⁵ An equally small number of controls would provide little ability to find associations. Increasing the number of controls up to a ratio of about 4/1 improves the power of the study. This rise is not linear, however. Beyond a ratio of about 4/1,³⁶ little power improvement results from increasing the number of controls.³⁷ Boosting the ratio of controls to cases affects the confidence interval (the precision of the results) but does not address bias.

What to look for in controls

The validity of case-control studies depends on selection of appropriate control groups. Choosing controls might seem deceptively simple but it can be treacherous. Controls should reflect the background frequency of the exposure in the population. Hence, they should be similar in all important respects to cases, except that they do not have the disorder in question. Their selection must be independent of exposure.

When the study group of potential controls is known, a good approach is to take all, or, if not feasible, a random sample of them. When the group of potential controls is unknown, choosing controls gets tough. Generally, we use individuals chosen from the same time and place as cases. Look for one good control group; if the appropriateness of a control group is uncertain, sometimes a second control group is added. If the number of cases is small, having up to four times as many controls improves study power. However, this strategy does not improve validity.

Use of inappropriate control groups generally leads to both wrong conclusions and potential medical harm.^{13,25,38,39} Readers of case-control reports need to think carefully about the characteristics of the controls. The results hang in the balance.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

We thank Willard Cates and David L Sackett for their helpful comments on an earlier version of this report.

References

- 1 Grimes DA. Clinical research in ancient Babylon: methodologic insights from the book of Daniel. *Obstet Gynecol* 1995; **86**: 1031–34.
- 2 Feinstein AR. Clinical epidemiology: the architecture of clinical research. Philadelphia: WB Saunders, 1985.
- 3 Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet* 2002; **359**: 431–34.
- 4 Sackett DL. Bias in analytic research. *J Chronic Dis* 1979; **32**: 51–63.
- 5 Marbach JJ, Schwartz S, Link BG. The control group conundrum in chronic pain case/control studies. *Clin J Pain* 1992; **8**: 39–43.
- 6 Rothman KJ. Modern epidemiology. Boston: Little, Brown and Company, 1986.
- 7 Schlesselman JJ. Case-control studies: design, conduct, analysis. New York: Oxford University Press, 1982.
- 8 Chan AT. Aspirin, non-steroidal anti-inflammatory drugs and colorectal neoplasia: future challenges in chemoprevention. *Cancer Causes Control* 2003; **14**: 413–18.
- 9 Baron JA. Epidemiology of non-steroidal anti-inflammatory drugs and cancer. *Prog Exp Tumor Res* 2003; **37**: 1–24.
- 10 Zondervan KT, Cardon LR, Kennedy SH. What makes a good case-control study? Design issues for complex traits such as endometriosis. *Hum Reprod* 2002; **17**: 1415–23.
- 11 Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies, I: principles. *Am J Epidemiol* 1992; **135**: 1019–28.
- 12 Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. *Lancet* 2002; **359**: 341–45.
- 13 Moss AR, Osmond D, Bacchetti P, Chermann JC, Barre-Sinoussi F, Carlson J. Risk factors for AIDS and HIV seropositivity in homosexual men. *Am J Epidemiol* 1987; **125**: 1035–47.
- 14 Perillo MG. Choice of controls in case-control studies. *J Manipulative Physiol Ther* 1993; **16**: 578–85.
- 15 Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies, II: types of controls. *Am J Epidemiol* 1992; **135**: 1029–41.
- 16 Hartge P, Brinton LA, Rosenthal JF, Cahill JJ, Hoover RN, Waksberg J. Random digit dialing in selecting a population-based control group. *Am J Epidemiol* 1984; **120**: 825–33.
- 17 Olson SH, Kelsey JL, Pearson TA, Levin B. Evaluation of random digit dialing as a method of control selection in case-control studies. *Am J Epidemiol* 1992; **135**: 210–22.
- 18 Voigt LF, Koepsell TD, Daling JR. Characteristics of telephone survey respondents according to willingness to participate. *Am J Epidemiol* 2003; **157**: 66–73.
- 19 Funkhouser E, Macaluso M, Wang X. Alternative strategies for selecting population controls: comparison of random digit dialing and targeted telephone calls. *Ann Epidemiol* 2000; **10**: 59–67.
- 20 Vernick LJ, Vernick SL, Kuller LH. Selection of neighborhood controls: logistics and fieldwork. *J Chronic Dis* 1984; **37**: 177–82.
- 21 Ryu JE, Thompson CJ, Crouse JR III. Selection of neighborhood controls for a study of coronary artery disease. *Am J Epidemiol* 1989; **129**: 407–14.
- 22 Rooks JB, Ory HW, Ishak KG, et al. Epidemiology of hepatocellular adenoma: the role of oral contraceptive use. *JAMA* 1979; **242**: 644–48.
- 23 Grimes DA, LeBolt SA, Grimes KR, Wingo PA. Systemic lupus erythematosus and reproductive function: a case-control study. *Am J Obstet Gynecol* 1985; **153**: 179–86.
- 24 Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002; **359**: 248–52.
- 25 Grimes DA. Intrauterine devices and pelvic inflammatory disease: recent developments. *Contraception* 1987; **36**: 97–109.
- 26 West DW, Schuman KL, Lyon JL, Robison LM, Allred R. Differences in risk estimations from a hospital and a population-based case-control study. *Int J Epidemiol* 1984; **13**: 235–39.
- 27 Infante-Rivard C. Hospital or population controls for case-control studies of severe childhood diseases? *Am J Epidemiol* 2003; **157**: 176–82.
- 28 Olson SH, Kelsey JL, Pearson TA, Levin B. Characteristics of a hypothetical group of hospital controls for a case-control study. *Am J Epidemiol* 1994; **139**: 302–11.
- 29 Lopes CS, Rodrigues LC, Sichieri R. The lack of selection bias in a snowball sampled case-control study on drug abuse. *Int J Epidemiol* 1996; **25**: 1267–70.
- 30 Goldstein AM, Hodge SE, Haile RW. Selection bias in case-control studies using relatives as the controls. *Int J Epidemiol* 1989; **18**: 985–89.
- 31 Ibrahim MA, Spitzer WO. The case control study: the problem and the prospect. *J Chronic Dis* 1979; **32**: 139–44.
- 32 Hulka BS, Fowler WC Jr, Kaufman DG, et al. Estrogen and endometrial cancer: cases and two control groups from North Carolina. *Am J Obstet Gynecol* 1980; **137**: 92–101.
- 33 Moritz DJ, Kelsey JL, Grisso JA. Hospital controls versus community controls: differences in inferences regarding risk factors for hip fracture. *Am J Epidemiol* 1997; **145**: 653–60.
- 34 Schulz KF, Grimes DA. Unequal group sizes in randomised trials: guarding against guessing. *Lancet* 2002; **359**: 966–70.
- 35 Henderson BE, Preston-Martin S, Edmondson HA, Peters RL, Pike MC. Hepatocellular carcinoma and oral contraceptives. *Br J Cancer* 1983; **48**: 437–40.
- 36 Kearney DJ, Crump C, Maynard C, Boyko EJ. A case-control study of endoscopy and mortality from adenocarcinoma of the esophagus or gastric cardia in persons with GERD. *Gastrointest Endosc* 2003; **57**: 823–29.
- 37 Wiebe DJ. Firearms in US homes as a risk factor for unintentional gunshot fatality. *Accid Anal Prev* 2003; **35**: 711–16.
- 38 Rossouw JE. Estrogens for prevention of coronary heart disease: putting the brakes on the bandwagon. *Circulation* 1996; **94**: 2982–85.
- 39 Leventhal JM. Risk factors for child abuse: methodologic standards in case-control studies. *Pediatrics* 1981; **68**: 684–90.