BMJ

BMJ 2014;348:g1072 doi: 10.1136/bmj.g1072 (Published 24 January 2014)

ENDGAMES

STATISTICAL QUESTION

Retrospective cohort studies: advantages and disadvantages

Philip Sedgwick *reader in medical statistics and medical education*

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

Researchers investigated whether differences exist between the sexes in the risk of ischaemic stroke in patients with atrial fibrillation.[1] A nationwide retrospective cohort study design was used. Data were taken from the Swedish national discharge register. Participants were 100 802 patients with a first diagnosis of atrial fibrillation between 1 July 2005 and 31 December 2008, with a total follow-up of 139 504 years at risk (median 1.2). Information about drug treatment was taken from the Swedish drug register. Patients were excluded if at baseline they were prescribed warfarin, had mitral stenosis, or had previous valvular surgery. Patients who died less than 14 days from baseline were also excluded.

The primary outcome measure was the occurrence of ischaemic stroke. It was reported that ischaemic strokes were more common in women than in men (6.2% *v* 4.2% per year; P<0.0001). The unadjusted hazard ratio of ischaemic stroke for women compared with men was 1.47 (95% confidence 1.40 to 1.54). After adjustment for 35 co-risk factors for stroke, an increased risk of stroke in women remained (1.18, 1.12 to 1.24). It was concluded that women with atrial fibrillation have a moderately increased risk of stroke compared with men, and thus, female sex should be considered when making decisions about anticoagulation treatment.

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

a) Patient data were collected retrospectively

b) Selection bias was minimised

c) Recall bias was minimised

d) It was possible to estimate the population at risk

e) Causality could be inferred from the association between female sex and ischaemic stroke in patients with atrial fibrillation

Answers

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

The aim of the study was to investigate whether differences exist between the sexes in the risk of ischaemic stroke in patients with atrial fibrillation. A retrospective cohort study design was used. Retrospective cohorts are observational in design and sometimes referred to as historic cohorts. Data were taken from the Swedish national discharge register, a database of the records of all hospital admissions and visits to hospital outpatient clinics since 1987 for patients with a Swedish civic registration. Information about drug treatment was taken from the Swedish drug register. Patients were included in the cohort study if they had a diagnosis of atrial fibrillation between 1 July 2005 and 31 December 2008. The cohort consisted of 100 802 patients, of whom 50 667 were women (50.3%). The example above is typical of a retrospective cohort study, where health records that have already been collected and stored in an electronic database are used to explore the association between one or more risk factors and a disease or condition.

The Swedish national discharge register had not been initially constructed with the aim of identifying a cohort to investigate the association between biological sex and stroke in patients with a diagnosis of atrial fibrillation. The above cohort study is described as retrospective because it involved looking back at events that had already taken place and been recorded in the register. Those patients diagnosed with atrial fibrillation between 1 July 2005 and 31 December 2008 were identified as the cohort. In effect, their experience was reconstructed as if they had been followed prospectively, particularly the subsequent occurrence of ischaemic stroke. Despite the above study being labelled as retrospective, the patient data would have been collected prospectively (*a* is false). The register would have been updated regularly as hospital admissions and visits to hospital outpatient clinics occurred.

Many of the advantages and disadvantages of retrospective cohort studies are similar to those for prospective cohort study designs. Prospective cohort studies have been described in a previous question.[2] As described above, retrospective cohort studies are typically constructed from databases of healthcare records that have already been collected. In contrast, a prospective design typically involves identifying a unique cohort that is followed prospectively, with the aim of investigating the association between one or more risk factors and a disease or

p.sedgwick@sgul.ac.uk

condition. However, an advantage for both study designs is that exposure to risk factors is recorded before the occurrence of the outcome. This is important because it allows the temporal sequence of risk factors and outcomes to be assessed. In particular, in the example above it permitted the epidemiology of ischaemic stroke to be studied.

Selection bias would have occurred if the cohort selected from the Swedish national discharge register was not representative of all possible patients with atrial fibrillation in the population. Confusion often exists as to what is meant by the "population" in statistics, probably because it has a different meaning from its general everyday one, where it is used in a geographical sense. Statistically, the population is typically regarded as an infinite group of people. The cohort study was a nationwide one—sometimes referred to as population based—and all patients with diagnosed atrial fibrillation in the population of Sweden between 1 July 2005 and 31 December 2008 were included. Therefore, the cohort members should have been representative of the population of all patients with atrial fibrillation. Hence, selection bias would have been minimised (*b* is true).

Recall bias, described in a previous question,[3] is typically associated with case-control studies that are retrospective in design. It is the systematic difference between those with a diagnosed disease or condition (the cases) and otherwise healthy people (the controls) in the accuracy of reported information about past exposure to risk factors. More generally, recall bias will originate if there are selective preconceptions between groups of patients about the association between the risk factor(s) and the outcome or condition. The health records in the Swedish national discharge register were collected prospectively and the database was not established to study risk factors for ischaemic stroke in patients with atrial fibrillation. Therefore, there was no obvious reason why systematic differences would have existed between groups of patients in the cohort in the accuracy of reported information. Recall bias would therefore have been minimised (*c* is true).

Because the cohort in the study above was population based, it was representative of the population. It was therefore possible to estimate the population at risk (*d* is true). Estimating the population at risk has been described in a previous question.[4] Being able to estimate the population at risk is an advantage, not least because the risk or incidence of ischaemic stroke in the cohort (patients with atrial fibrillation) as a whole, and for men and women separately, can be used to estimate the risk in the population.

As described above, the Swedish national discharge register was not initially constructed with the aim of identifying a cohort to investigate the association between biological sex and stroke in patients with a diagnosis of atrial fibrillation. However, the use of records that had already been collected and stored in an electronic database meant that this retrospective cohort study was relatively cheap, quick, and easy to perform. This is particularly so compared with using a prospective cohort study design to investigate the association between biological sex and stroke. Nonetheless, a consequence of retrospective cohort studies using health records that have already been collected is that not all pertinent risk factors are likely to have been identified and subsequently recorded. A further disadvantage of retrospective cohort studies is that many different healthcare professionals will have been involved in patient care, so the measurement of risk factors and outcome(s) throughout the database would probably be less accurate and consistent than that achieved with a prospective cohort study design.

In the retrospective cohort study above, participants were patients with a diagnosis of atrial fibrillation between 1 July 2005 and 31 December 2008. Patients were followed until the occurrence of ischaemic stroke, which was diagnosed using the international classification of diseases, 10th revision. The period chosen for inclusion of patients from the Swedish national discharge register was relatively short, and the median follow-up was only 1.2 years. It was therefore possible to achieve a consistent diagnosis of the outcome of ischaemic stroke. However, some retrospective cohort studies have a substantial length of follow-up, and it may be difficult to ensure outcomes are measured consistently or using the same criteria. Furthermore, when cohort studies have a substantial length of follow-up, the association between the risk factor(s) and the outcome or condition may change with time.

The Swedish national discharge register contained records for a large number of hospital admissions and visits to hospital outpatient clinics. It is possible that not all records were complete. If patients who were part of the cohort identified in the above study had incomplete records, it may have biased any observed associations. This will be particularly so if data were not missing at random—that is, if the reason for missing data was related to the risk factor of biological sex or the outcome of ischaemic stroke in patients with atrial fibrillation.

As is typical of observational studies, only association and not causation can be inferred from the results of the above cohort study (*e* is false). This is because the observed association between female sex and ischaemic stroke may have been the result of confounding. In particular, it was not possible to measure and then control for, through statistical analysis, all factors that may have affected the outcome of ischaemic stroke. This is despite the hazard ratio for ischaemic stroke for women relative to men being adjusted for 35 potential co-risk factors—recording exposure to a wide range of risk factors is fairly typical in a cohort study. In contrast, experimental studies such as clinical trials use random allocation of participants to treatment groups to control for confounding at baseline. Nonetheless, the statistician Austin Bradford-Hill proposed criteria, which, if met, may allow causation to be inferred from an association between a risk factor and outcome in observational studies.[5] Further discussion of the criteria is beyond the scope of this article.

Competing interests: None declared.

1 Friberg L, Benson L, Rosenqvist M, Lip GYH. Assessment of female sex as a risk factor in atrial fibrillation in Sweden: nationwide retrospective cohort study. *BMJ* 2012;344:e3522.
2 Sedgwick P. Prospective cohort studies: advantages and disadvantages. *BMJ* 2013;347:f6726.
3 Sedgwick P. What is recall bias? *BMJ* 2012;344:e3519.
4 Sedgwick P. Estimating the population at risk. *BMJ* 2012;345:e6859.
5 Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295-300.

Cite this as: *BMJ* 2014;348:g1072