ENSURE - Educating students for developing high quality research skills

This material was realised with the EEA Financial Mechanism 2014-2021 financial support. Its content (text, photos, videos) does not reflect the official opinion of the Programme Operator, the National Contact Point and the Financial Mechanism Office. Responsibility for the information and views expressed therein lies entirely with the author(s).

UiT

NORGES
ARKTISKE
UNIVERSITET

23. October 2019

Case-control design

Finn Egil Skjeldestad
Institute of community medicine
UiT The arctic university of Norway
Tromsø
Norway

Case-control study

Definition

- A study that **compares patients who have a disease or outcome of interest** (cases) ***with patients who do not have the disease or outcome (controls)*****, and looks back retrospectively to compare how frequently the exposure to a risk factor is present in each group to determine the relationship between the risk factor and the disease.**

- Case control studies are **observational because no intervention** is attempted and no attempt is made to alter the course of the disease.

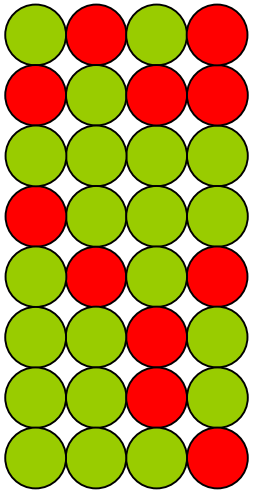
- The goal is to **retrospectively determine the exposure to the risk factor of interest** from each of the two groups of individuals: cases and controls.

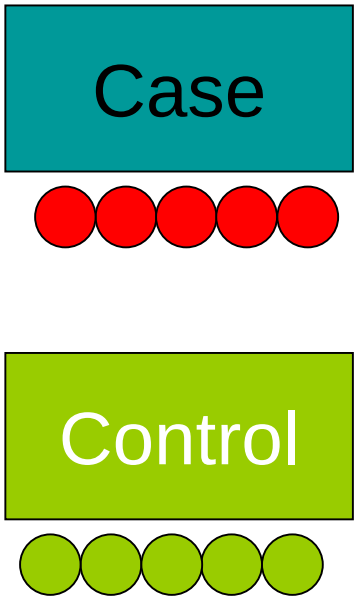
- These studies are designed to **estimate odds.**

Case-control design

Participants selected on disease status

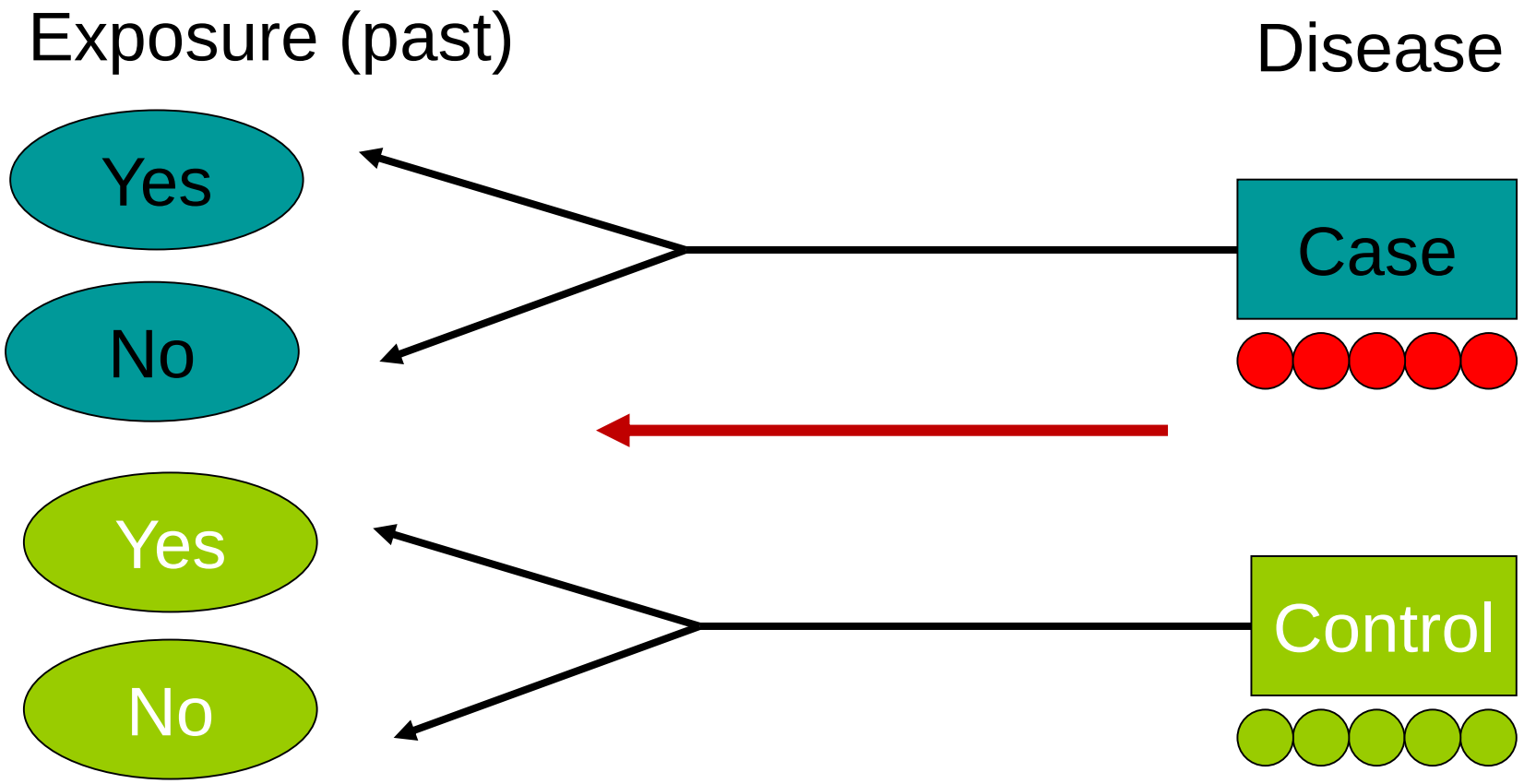
Population





Diseases No disease

Case-control design



Cases

Incident cases

- Comprise cases newly diagnosed during a defined time period
- The use of incident cases is considered as preferential, as the recall of past exposure(s) may be more accurate among newly diagnosed cases.
- **The temporal sequence of exposure and disease** is easier to assess among incident cases.

Prevalent cases

- Comprise individuals who have had the outcome under investigation for some time.
- The use of prevalent cases may give rise to recall bias as prevalent cases may be less likely to accurately report past exposures(s).
- The interpretation of results based on prevalent cases may prove more problematic, as it may be more difficult to ensure that reported events relate to a time before the development of disease rather than to the consequence of the disease process itself.

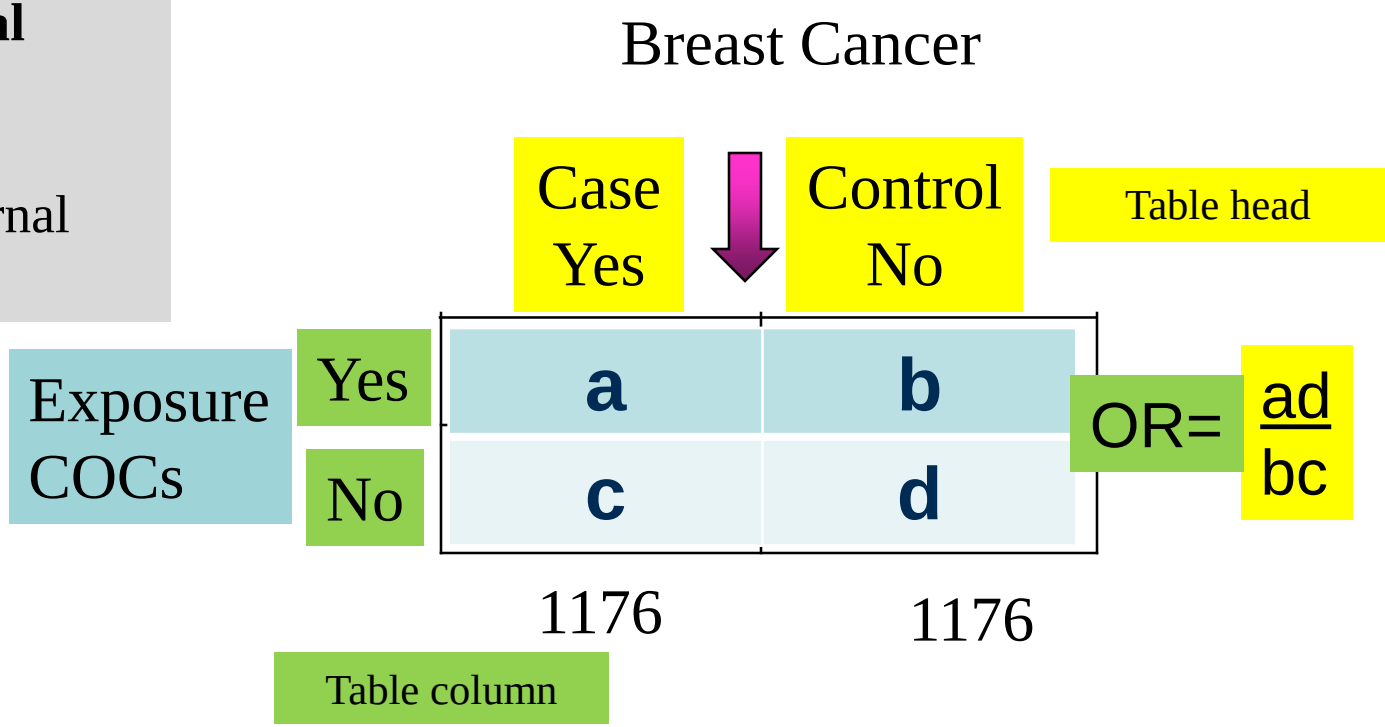
Selection of controls

Controls are used to estimate the prevalence of exposure in the population which gave rise to the cases.

- The ideal control group would comprise a random sample from the general population that gave rise to the cases.
- The goal is to select individuals in whom the distribution of exposure status would be the same as that of the cases in the absence of an exposure disease association. That is, if there is no true association between exposure and disease, the cases and controls should have the same distribution of exposure
- Controls should be selected to be a representative sample of the population which produced the cases.
  - If cases are selected from a defined population such as a GP register, then controls should comprise a sample from the same GP register.
  - If case are hospital based, it is common to recruit controls from the hospital population.
- Recruiting more than one control per case may improve the statistical power of the study, though including more than 4 controls per case is generally considered to be no more efficient.

Is risk of BC increased in current COCs users?

Recruitment - hospital
Women: 16-50 yrs.
Controls:
Patients, admitted internal med. or surgical wards



Odds ratio/ OR=(537*622)/(639*554) = 0.94

MATCHED CASE-CONTROL

Cases and controls are matched for

- Age
- Sex
- Residency
- ….etc.

When matching,
there is no difference in factors/exposures matched for

Nested CASE-CONTROL study

In the nested case-control study,

cases of a disease that occur in a defined cohort are identified and, for each, a specified number of matched controls is selected from among those in the cohort who have not developed the disease by the time of disease occurrence in the case.

The nested case-control design potentially offers impressive reductions in

- costs
- efforts of data collection, analysis compared with the full cohort approach, with relatively minor loss in statistical efficiency.
- The nested case-control design is particularly advantageous for studies of biologic precursors of disease (biobanks).

Data sources – case-control design

- Standardized questionnaires
- Interviews with the subject
- Interviews with spouse or other family members
- Medical records
- Employment records
- Pharmacy records (national/regional/local)
- Biological samples

Unsystematic data collection in favour of cases

Bias in case-control studies

- Selection bias (selection of controls)
- Recall bias (prevalent/incident disease)
- Interviewer/observer bias
  - recording of exposure information may vary depending on the investigator's knowledge of an individual's disease status

Case-control design

Advantages:

- Relatively cheap
- Relatively few people are investigated
- The result comes quickly
- Good for studying rare conditions or diseases
- Less time needed to conduct the study because the condition or disease has already occurred
- Lets you simultaneously look at multiple risk factors
- Useful as initial studies to establish an association
- Can answer questions that could not be answered through other study designs

Disadvantages:

- The information from the case may be affected by the disease
- Finding good controls can be difficult
- Cannot calculate incidence
- Retrospective studies have more problems with data quality because they rely on memory and people with a condition will be more motivated to recall risk factors (recall bias).
- Not good for evaluating diagnostic tests because its already clear that the cases have the condition and the controls do not