Lecture 9: Case Control Studies

Lecture prepared by Dr. Hailey Banack, PhD

Cohort Studies Review

- Advantages: measurement of exposure, outcome, covariates decided at baseline, can assess temporal ordering of exposure and disease

- Disadvantages: expensive, can be inefficient and time consuming

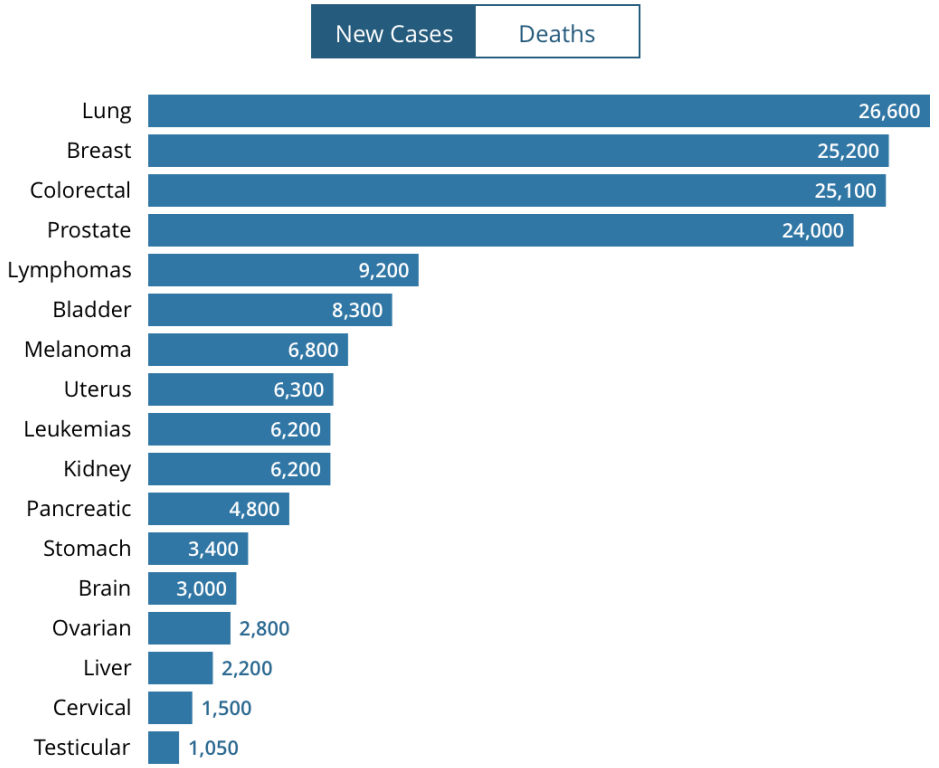
Will there be enough cases to study?

Say we're interested in studying whether occupational exposure to xray increases the risk of testicular cancer among male health care workers.

Very few cases of testicular cancer are diagnosed each year.

Sometimes it may be useful to start by gathering cases of disease.

Projected new cases and deaths from cancer in Canada for 2015



Statistics source: Canadian Cancer Society

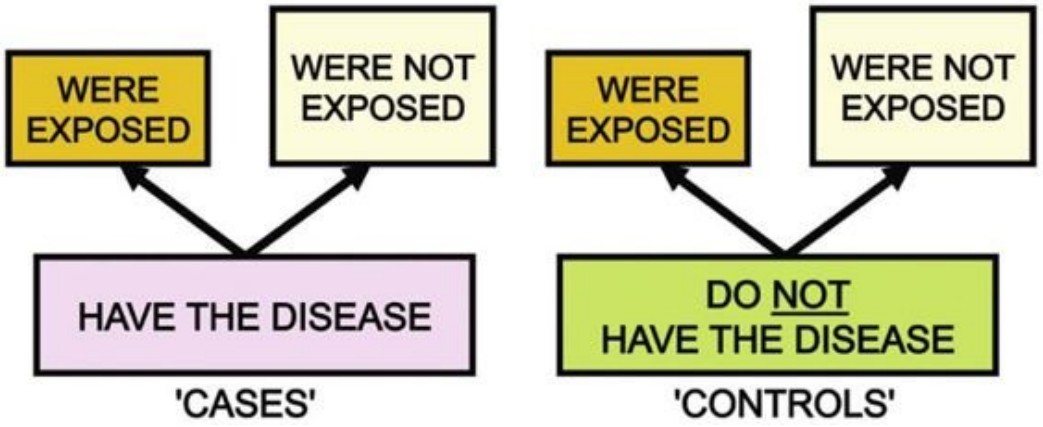
These 17 cancer types represent 84 per cent of the new cases expected to be diagnosed in Canada in 2015

Why a case-control design?

- Outcome being studied is rare
- Outcome has long induction period or long latent period
  - Etiologic relationship takes years to manifest
- Exposure data is difficult or expensive to obtain
  - Biologic/genetic measurements, record abstraction

Case-Control Design

- Compare cases (diseased) and controls (non-diseased) with respect to exposure level
- Contrast the odds of exposure among cases with odds of exposure among controls
- Cases and controls sampled from a “study base” (ideally the same study base)

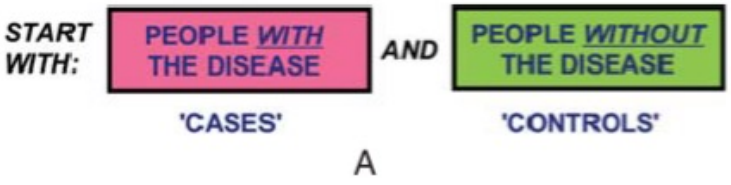


Basic case control design

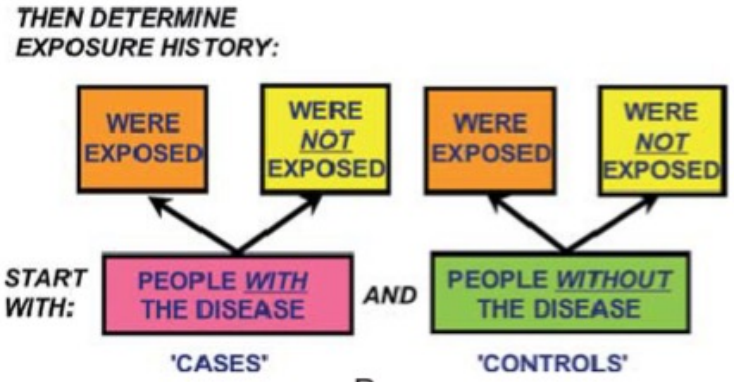
1. Start with cases of disease (D+/D-)

2. Determine exposure status (E+/E-)

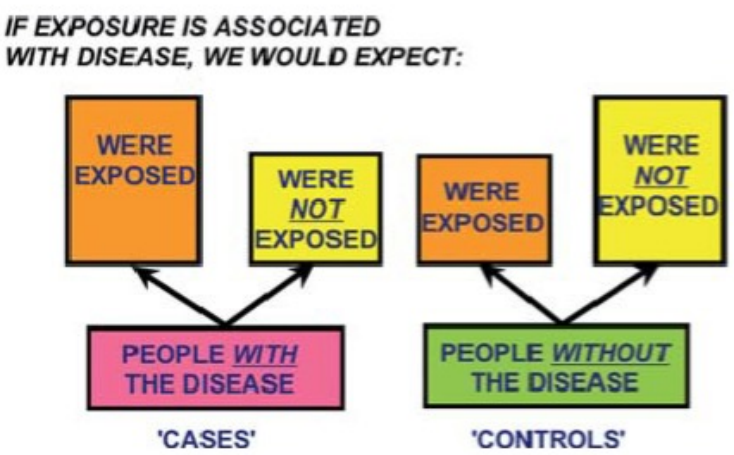
	First Select	
	Cases (With Disease)	***Controls (Without Disease)***
Then Measure Past Exposure		
Were exposed	a	b
Were not exposed	c	d
Total	$a + c$	$b + d$
Proportions exposed	$\frac{a}{a + c}$	$\frac{b}{b + d}$


START WITH:
PEOPLE WITH THE DISEASE
AND
PEOPLE WITHOUT THE DISEASE
'CASES'
'CONTROLS'

A


THEN DETERMINE EXPOSURE HISTORY:
WERE EXPOSED
WERE NOT EXPOSED
WERE EXPOSED
WERE NOT EXPOSED
START WITH:
PEOPLE WITH THE DISEASE
AND
PEOPLE WITHOUT THE DISEASE
'CASES'
'CONTROLS'

B


IF EXPOSURE IS ASSOCIATED WITH DISEASE, WE WOULD EXPECT:
WERE EXPOSED
WERE NOT EXPOSED
WERE EXPOSED
WERE NOT EXPOSED
PEOPLE WITH THE DISEASE
PEOPLE WITHOUT THE DISEASE
'CASES'
'CONTROLS'

C

Case control examples

Hypothetical Example of a Case-Control Study of Coronary Heart Disease and Cigarette Smoking

	CHD Cases	Controls
Smoke cigarettes	112	176
Do not smoke cigarettes	88	224
Total	200	400

Distribution of 1,357 Male Lung Cancer Patients and a Male Control Group According to Average Number of Cigarettes Smoked Daily Over the 10 Years Preceding Onset of the Current Illness

Average Daily Cigarettes	Lung Cancer Patients	Control Group
0	7	61
1–4	55	129
5–14	489	570
15–24	475	431
25–49	293	154
50+	38	12
Total	1,357	1,357

From Doll R, Hill AB: A study of the aetiology of carcinoma of the lung. BMJ 2:1271–1286, 1952.

Pros and Cons

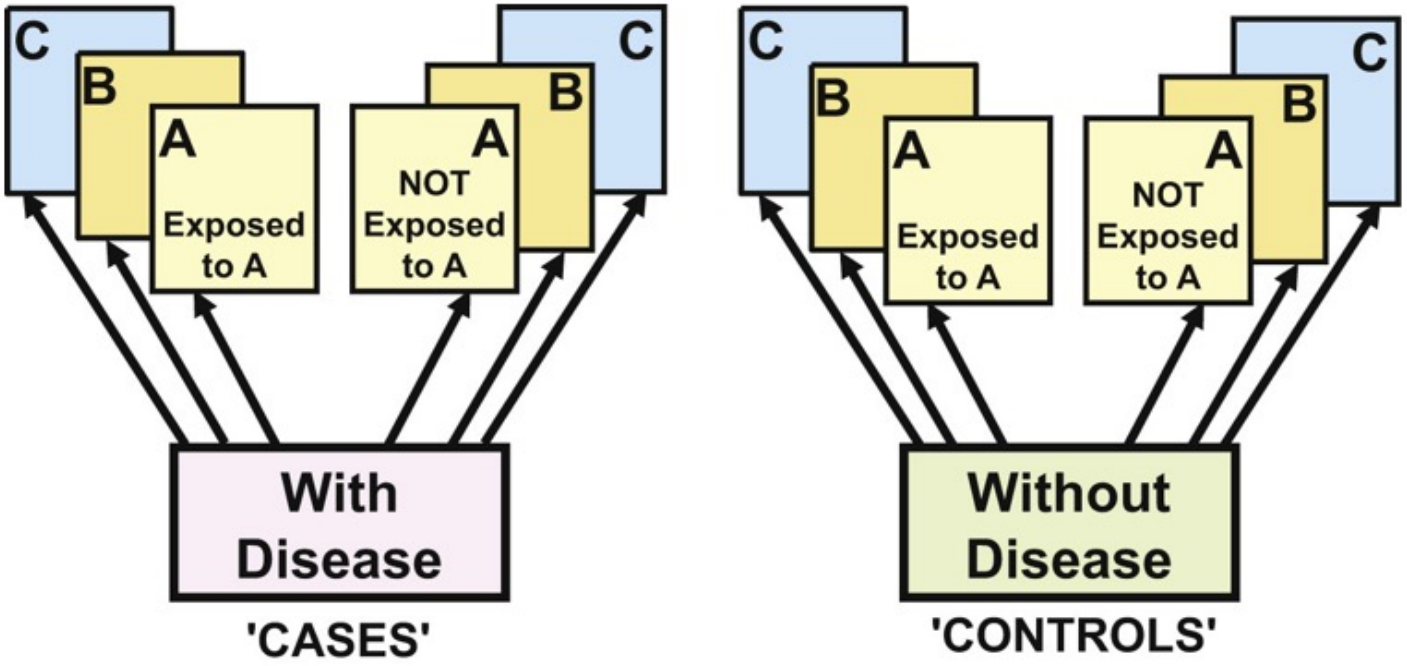
Pros

- More efficient than cohort studies
- Good for studying rare diseases or disease with long latency
- Cost saving
- Need for follow-up time is avoided

Cons

- Less efficient than cohort studies if exposure is rare
- Reduced precision compared with cohort studies
- Increased potential for bias

Case control studies: good for studying multiple exposures of interest



Study Base (Source Population)

- One of the most important and challenging aspects of designing a case control study is defining the study base
- Sample cases- from hospital, doctor's office, disease registry etc
  - Who are the comparable controls for these cases?
- Cases and controls must be drawn from the same 'study base'
- Controls must be members of the underlying source population/cohort from which the cases are drawn
- Want controls to represent the experience of the entire non-disease study base

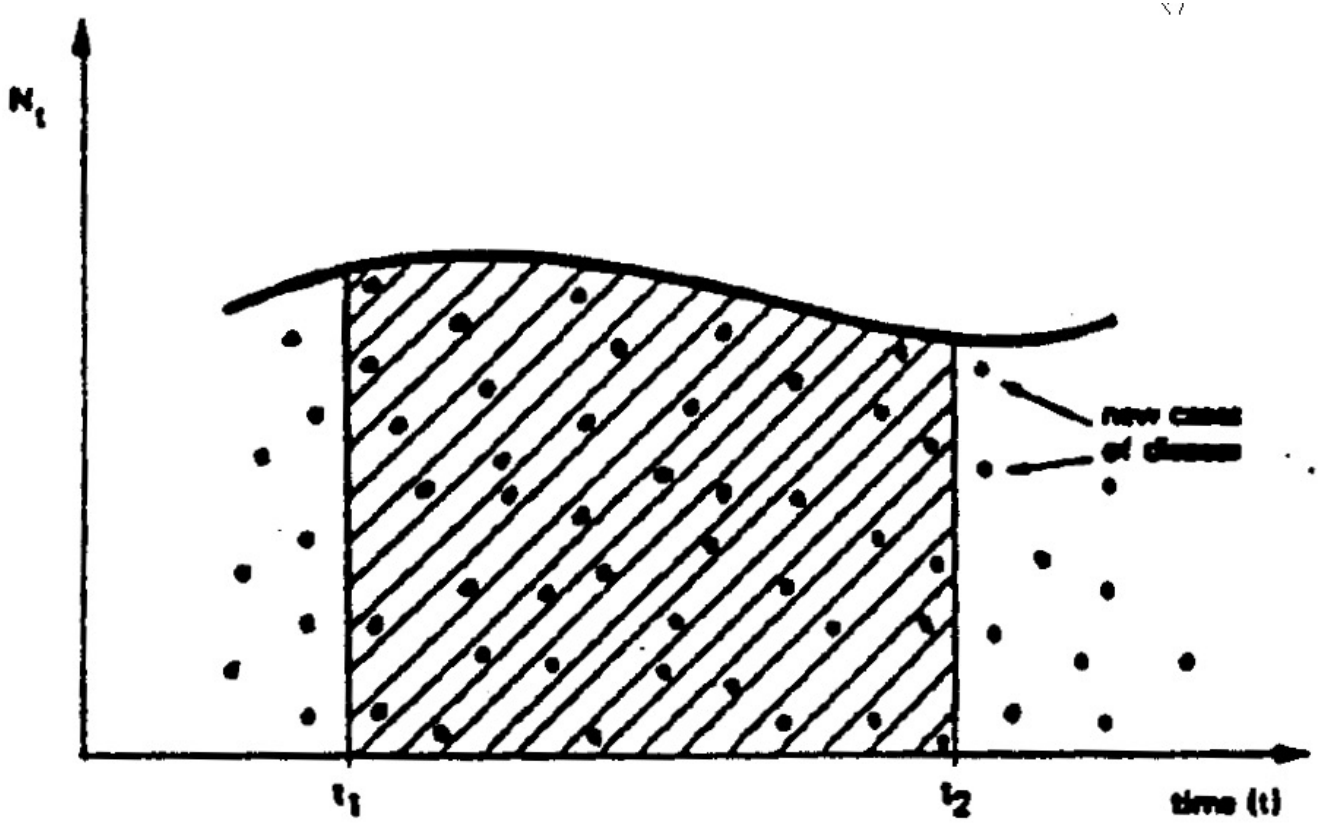


FIGURE 2 *Graphical illustration of the occurrence of new (incident) cases over time in a candidate population (of size N_t at time t)*

Morgenstern IJE 1980

COHORT STUDIES

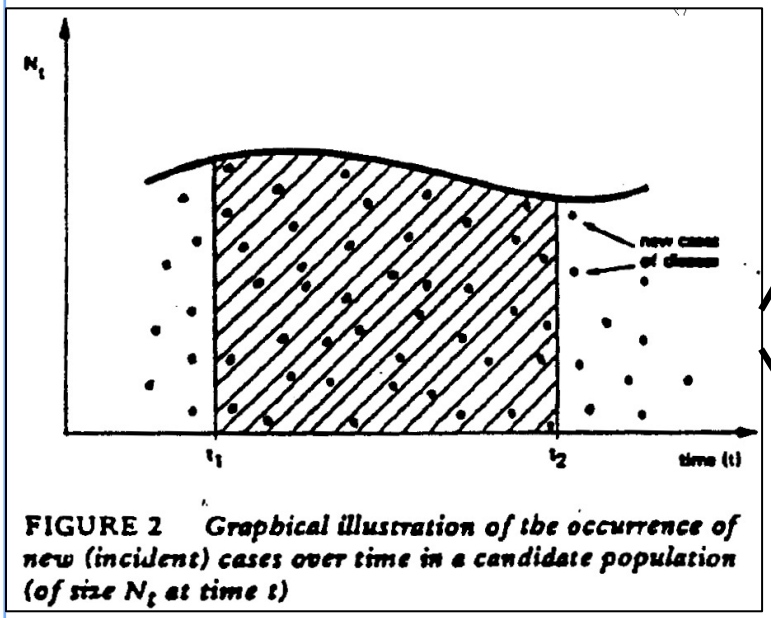
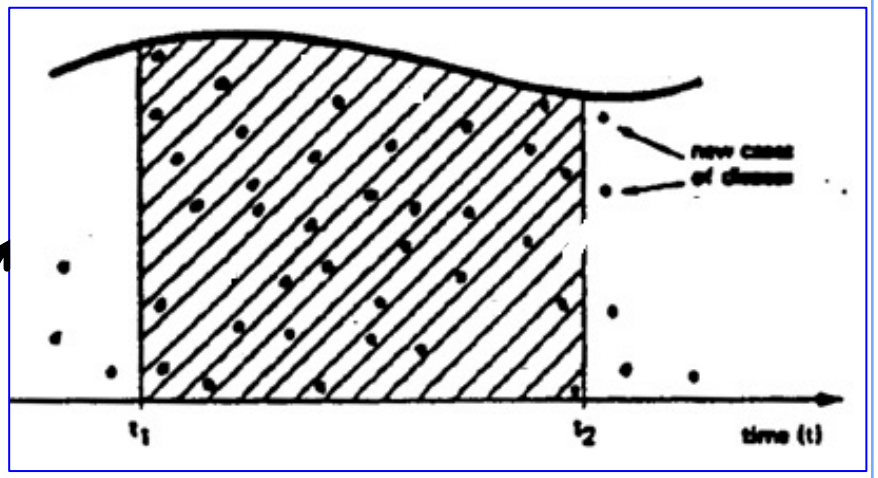


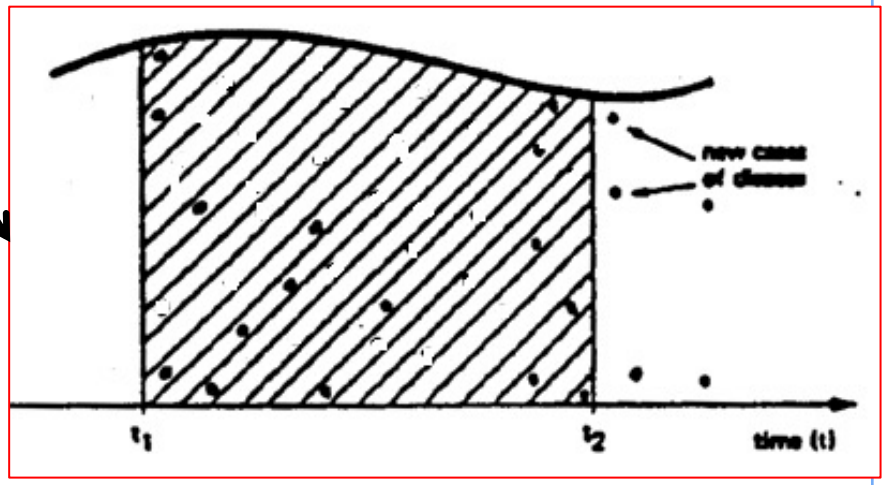
FIGURE 2 *Graphical illustration of the occurrence of new (incident) cases over time in a candidate population (of size N_t at time t)*

Total population

New cases/Exposed PT



New cases/Unexposed PT



CASE-CONTROL STUDIES

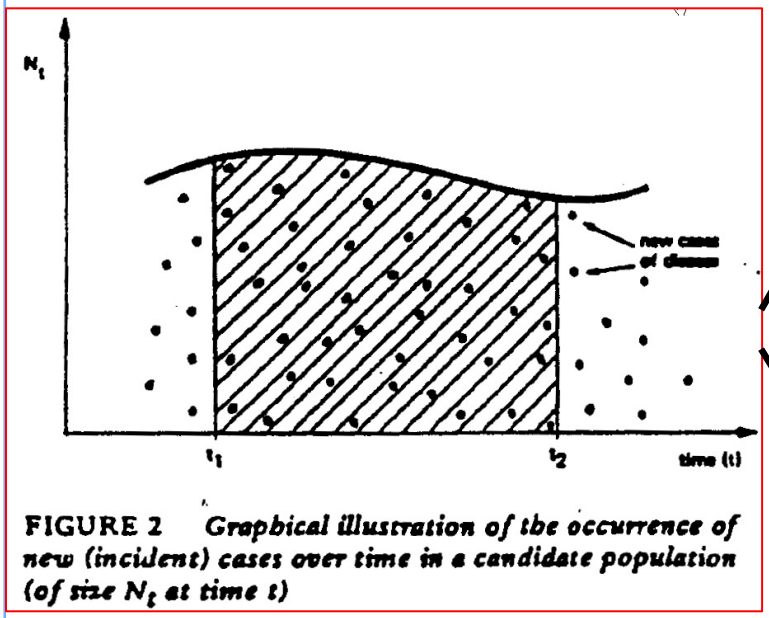
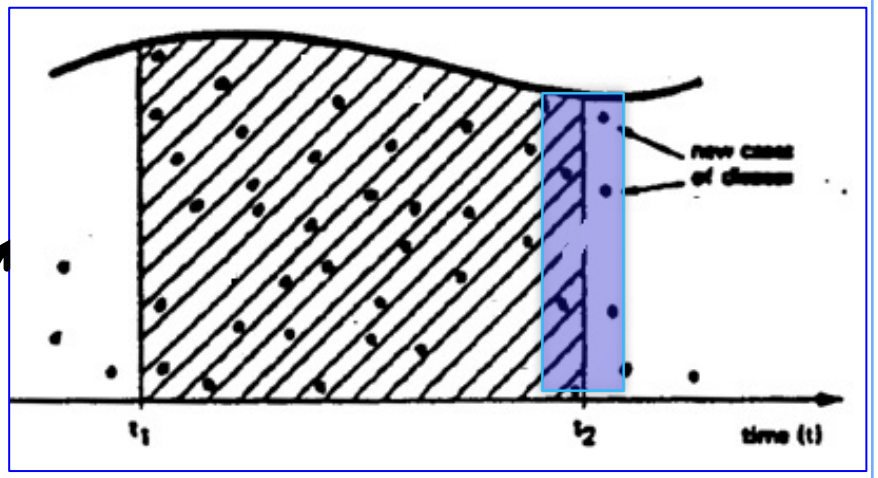


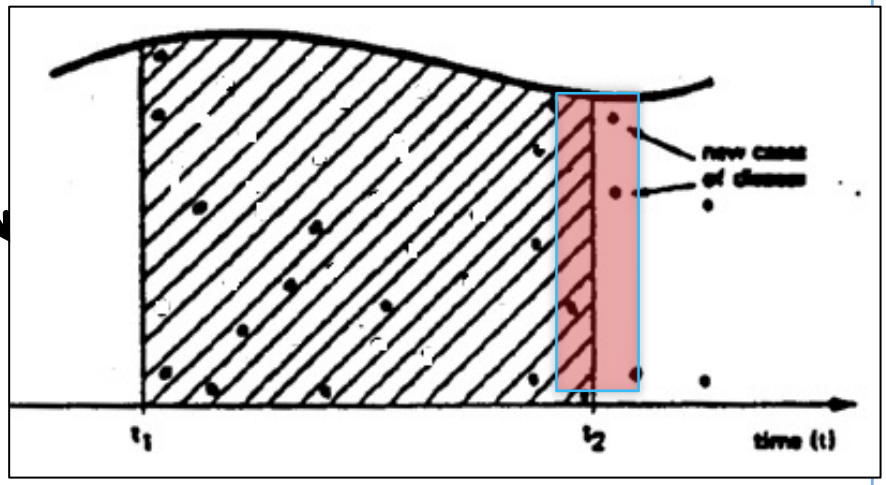
FIGURE 2 *Graphical illustration of the occurrence of new (incident) cases over time in a candidate population (of size N_t at time t)*

Total population

Obtain a representative sample of cases that occur in the study base (=case series)



Obtain a representative sample of the study base itself (=control series)



Conducting a traditional case-control study

1. Define the source population for the study
2. Define the exposures and outcome of interest
3. Identify cases (=D+) and determine exposure status
4. Sample controls from the source population and determine exposure status
   - The purpose of the control group is to represent the distribution of exposure in the source population
5. Calculate and interpret the odds ratio

Types of Case Controls Studies

In a defined cohort

1. Cumulative case control (traditional case control design)

Controls = those in the study base who do not experience the outcome

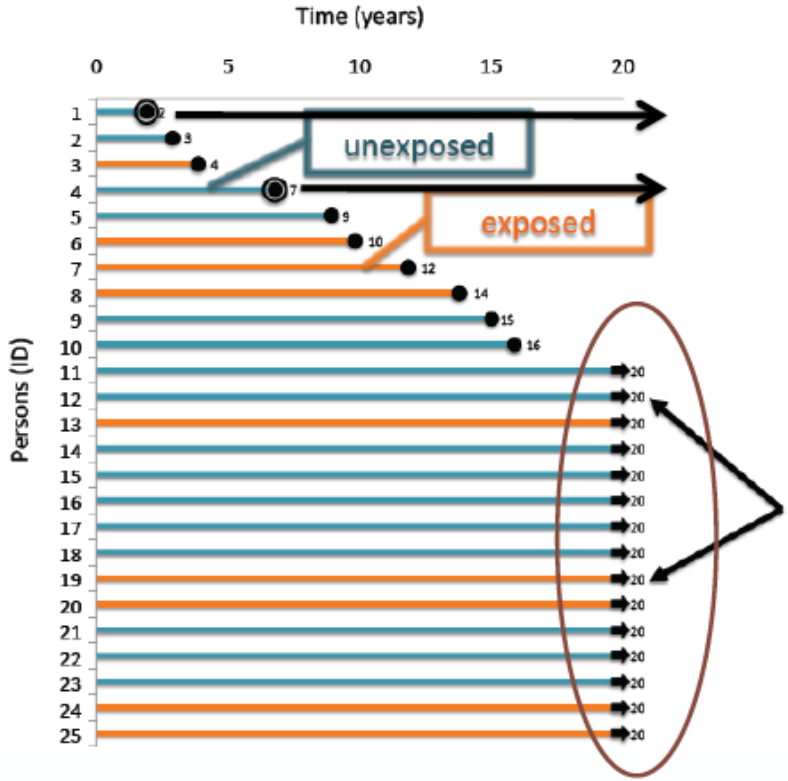
2. Case-cohort design

Controls sampled from the entire cohort at baseline (=start of follow-up period)

3. Nested (density) case control design

Controls selected throughout the course of the study, from individuals at risk (in the risk set) each time a case is diagnosed

Traditional case control study



Controls are selected from those in the population who do not have the outcome (disease) of interest

Odds Ratios

	Cases	Controls
Exposure +	a	b
Exposure -	c	d
Proportion Exposed	a/(a+c)	b/(b+d)

Odds Ratio: *[a * d] / [b * c]*

Odds Ratio in Cohort Study

	Develop Disease	Do Not Develop Disease
Exposed	a	b
Not Exposed	c	d

$$OR = \frac{\text{odds that an exposed person develops disease}}{\text{odds that a non-exposed person develops disease}}$$

$$= \frac{a/b}{c/d}$$

$$= \frac{ad}{bc}$$

A

Odds Ratio in Case Control Study

	CASES (with disease)	CONTROLS (without disease)
History of exposure	a	b
No history of exposure	c	d

$$OR = \frac{\text{odds that a case was exposed}}{\text{odds that a control was exposed}}$$

$$= \frac{a/c}{b/d}$$

$$= \frac{ad}{bc}$$

B

Odds Ratios with Multiple Exposure Categories

Average Daily Cigarettes	Lung Cancer Patients	Control Group
0	7	61
1–4	55	129
5–14	489	570
15–24	475	431
25–49	293	154
50+	38	12
Total	1,357	1,357

Distribution of 1,357 Male Lung Cancer Patients and a Male Control Group According to Average Number of Cigarettes Smoked Daily Over the 10 Years Preceding Onset of the Current Illness

2x2 Table comparing 1-4 daily cigarettes to 0 daily cigarettes		
	Cases	**Controls**
E+	55	129
E-	7	61

2x2 Table comparing 25-49 daily cigarettes to 0 daily cigarettes		
	Cases	**Controls**
E+	293	154
E-	7	61

Case–Control Study of Blood Lead Levels and Attention Deficit Hyperactivity Disorder in Chinese Children

Hui-Li Wang,[1] Xiang-Tao Chen,[1,2] Bin Yang,[3] Fang-Li Ma,[4] Shu Wang,[1] Ming-Liang Tang,[1] Ming-Gao Hao,[5] and Di-Yun Ruan[1]

Background: Attention deficit/hyperactivity disorder (ADHD) and lead exposure are high-prevalence conditions among children.

Objective: Our goal was to investigate the association between ADHD and blood lead levels (BLLs) in Chinese children, adjusting for known ADHD risk factors and potential confounding variables.

Methods: We conducted a pair-matching case–control study with 630 ADHD cases and 630 non-ADHD controls 4–12 years of age, matched on the same age, sex, and socioeconomic status. The case and control children were systematically evaluated via structured diagnostic interviews, including caregiver interviews, based on the *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., revised criteria (DSM-IV-R). We evaluated the association between BLLs and ADHD using the Pearson chi-square test for categorical variables and the Student *t*-test for continuous data. We then performed conditional multiple variables logistic regression analyses with backward stepwise selection to predict risk factors for ADHD.

Results: There was a significant difference in BLLs between ADHD cases and controls. ADHD cases were more likely to have been exposed to lead during childhood than the non-ADHD control subjects, with adjustment for other known risk factors [children with BLLs ≥ 10 µg/dL vs. ≤ 5 µg/dL; OR = 6.0; 95% confidence interval (CI) = 4.10–8.77, $p < 0.01$; 5–10 µg/dL vs. ≤ 5 µg/dL, OR = 4.9; 95% CI = 3.47–6.98, $p < 0.01$]. These results were not modified by age and sex variables.

Conclusions: This was the largest sample size case–control study to date to study the association between BLLs and ADHD in Chinese children. ADHD may be an additional deleterious outcome of lead exposure during childhood, even when BLLs are < 10 µg/dL.

Key words: attention deficit hyperactivity disorder, blood lead levels, case–control study. *Environ Health Perspect* 116:1401–1406 (2008). doi:10.1289/ehp.11400 available via *http://dx.doi.org/* [Online 5 June 2008]

ADHD cases were consecutively recruited from children coming for initial or follow-up assessment from October 2003-August 2007 in two pediatric clinics

Non-ADHD controls were randomly selected from computerized lists of outpatients admitted for acute respiratory infection at the same two clinics during the same time period

Case control studies in a defined cohort

-Taking advantage of the best features of cohort and case-control designs

-Case control study that takes place within a defined cohort

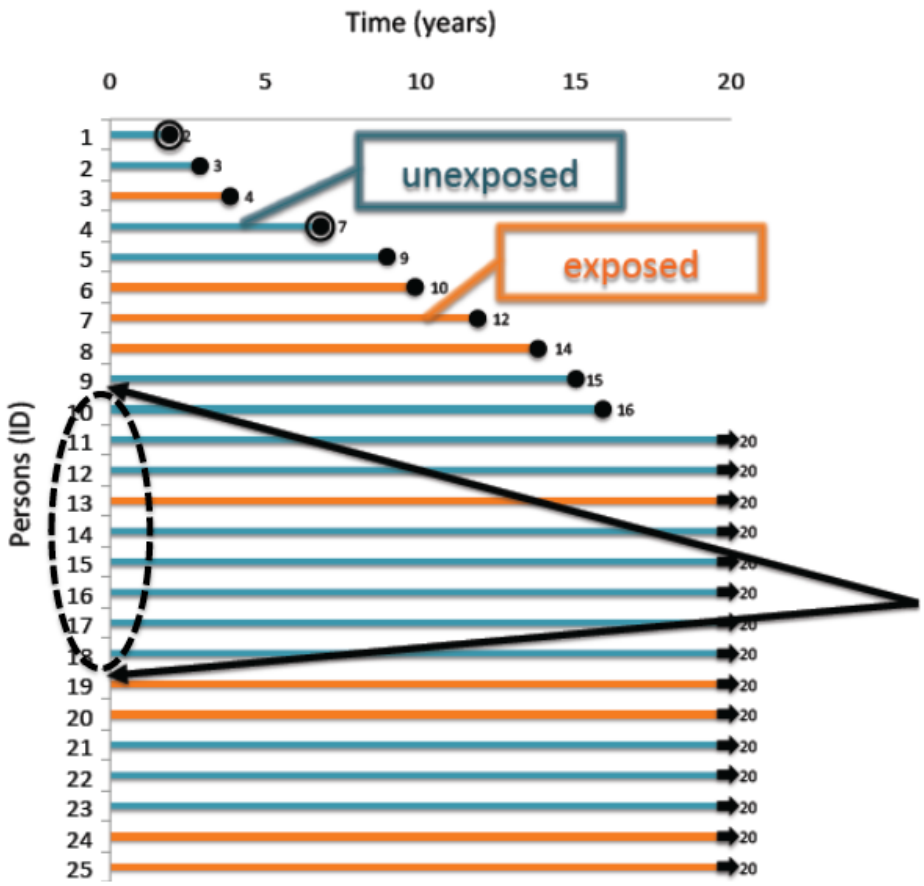
-Case cohort study

-Nested case control study

-Overall idea:

- Assemble cohort
- Follow over time
- Some individuals will develop disease (cases)
- Take a sample of those individuals who do not (controls)

Case Cohort Design



- Controls are selected from the entire source population (those at risk at the beginning of follow-up, including those who may experience the event over follow-up)
- Every person in the cohort has an equal chance of being included in the study as a control, regardless of person time in the cohort or whether they developed the disease

Are retinol, vitamin C, vitamin E, folate and carotenoids intake associated with bladder cancer risk? Results from the Netherlands Cohort Study

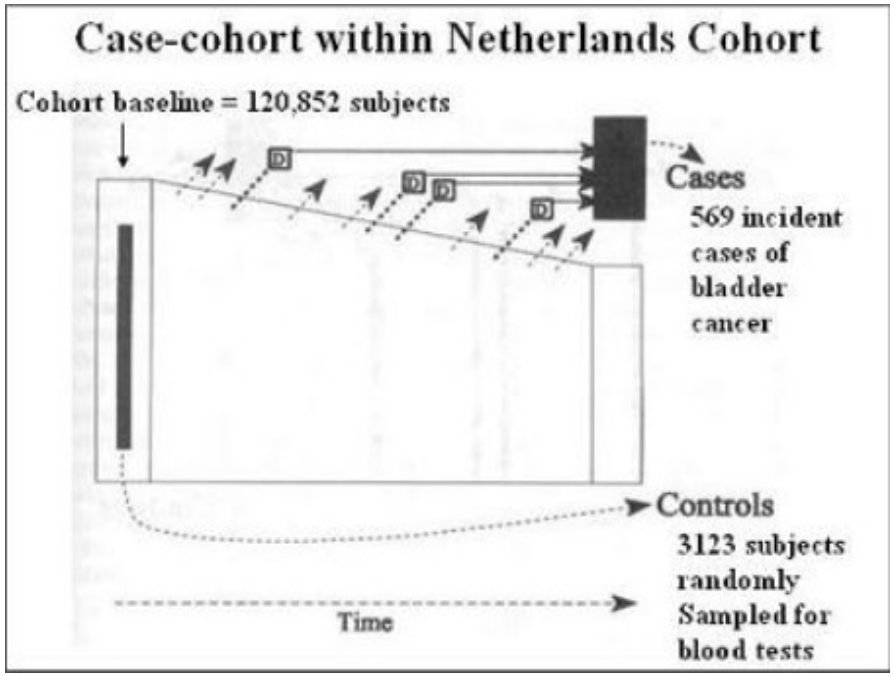
MPA Zeegers[1], RA Goldbohm[2] and PA van den Brandt[1]

[1]Department of Epidemiology, Maastricht University, Maastricht; and [2]Department of Nutritional Epidemiology, TNO Nutrition, and Food Research, Zeist, The Netherlands

A large cohort study in the Netherlands of 120,852 people followed since 1986

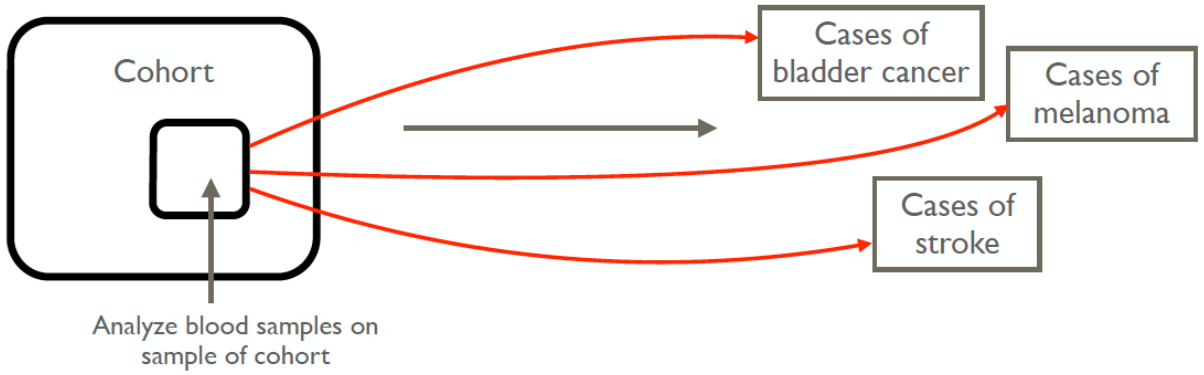
All cohort members completed baseline questionnaires and provided a blood sample

It would have been very expensive and inefficient to do lab testing on blood samples for the entire cohort, so a random sample of 3123 individuals was chosen from the baseline cohort

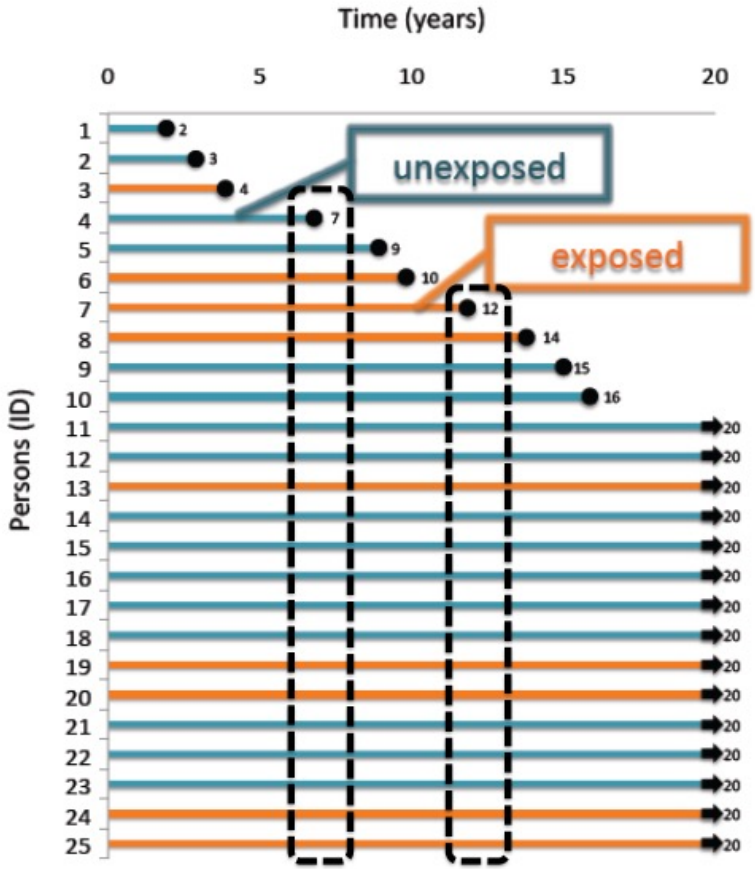


Advantage of a case-cohort design

- The same control series can be used in several studies



Nested Case Control studies



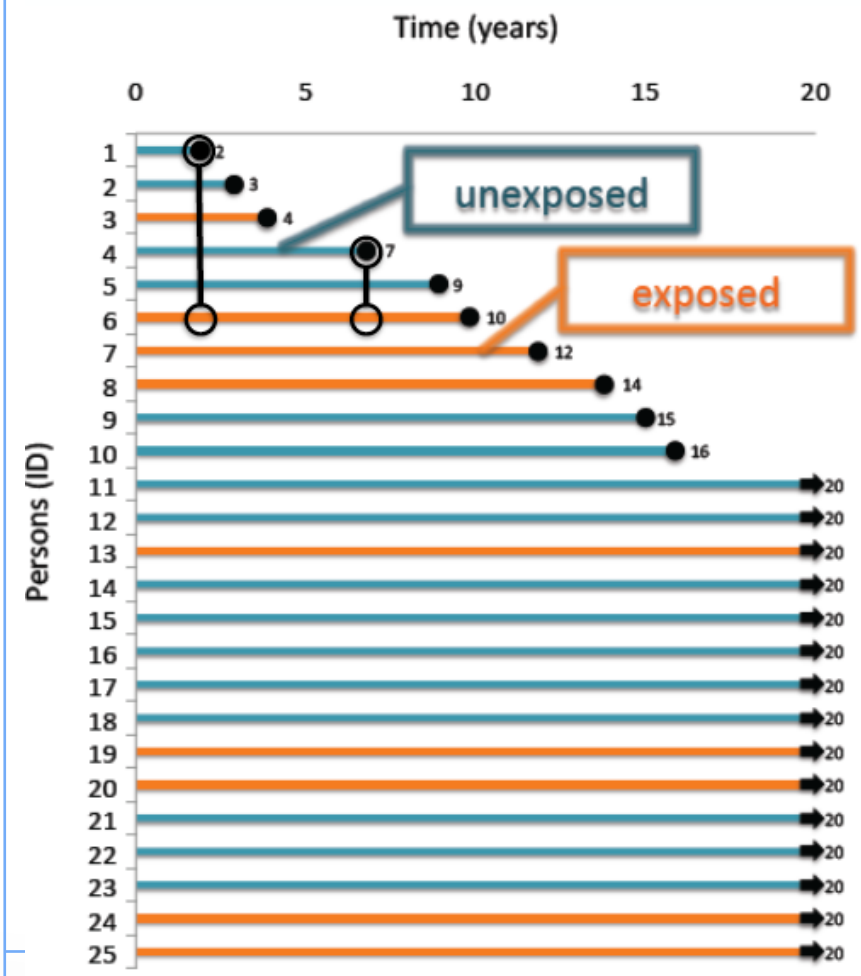
-Controls selected from individuals in the 'risk set' at the time a case occurs

-A person who has already developed the disease (=case) is NOT eligible to be sampled as a control

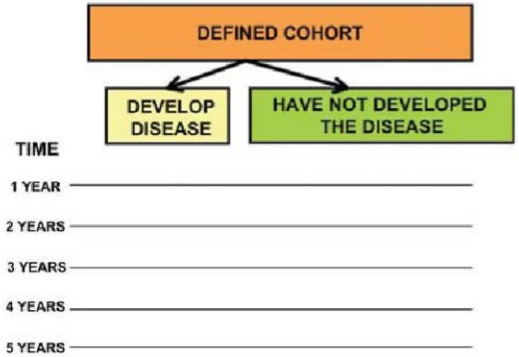
-However, controls are able to become cases if they later develop the disease

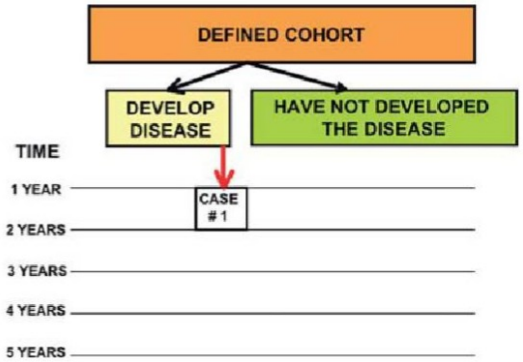
- An individual could be a control and then selected as a case at a later time point

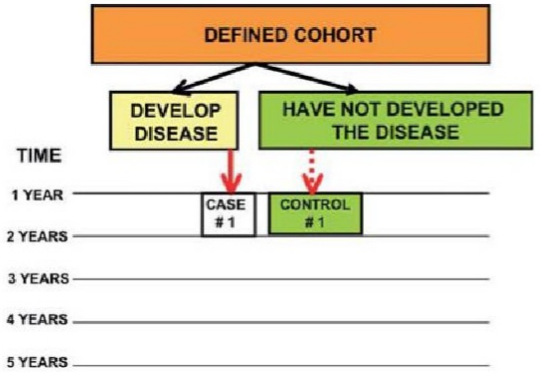
Nested Case Control studies

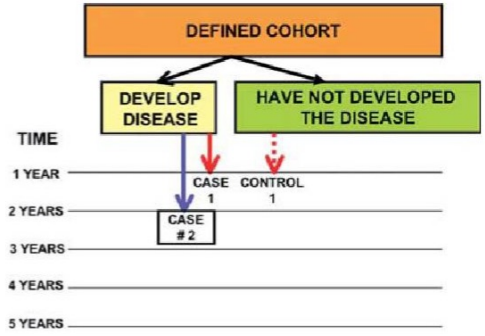


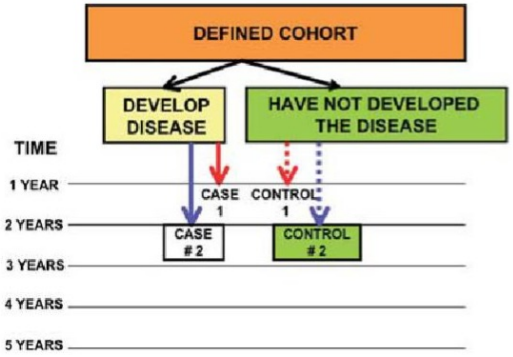
- A person selected as a control is still eligible to be selected again as a control as long as they remain at risk for disease
- The same person may serve as a control for multiple cases

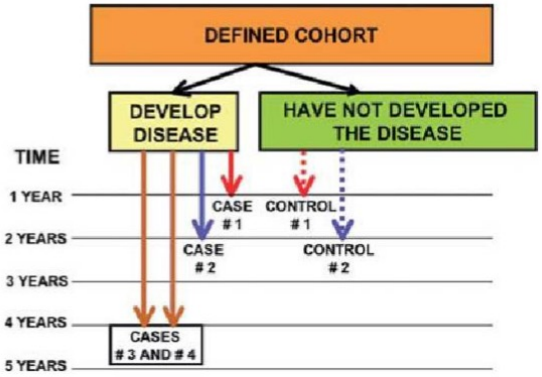

DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS

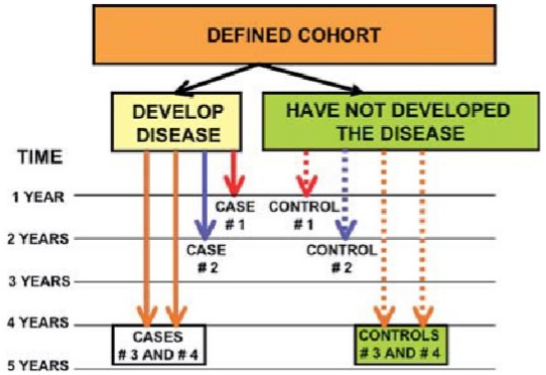

DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE #1

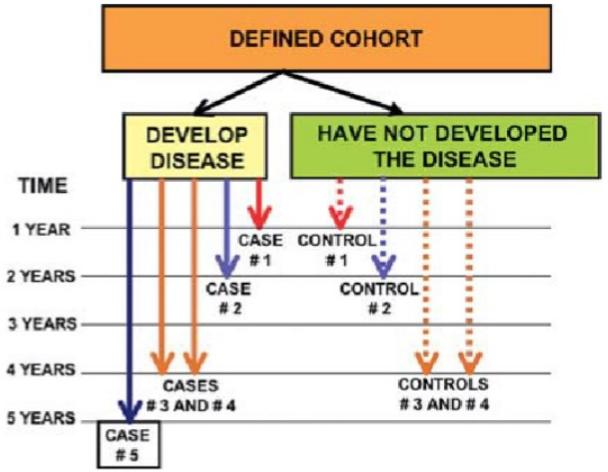

DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE #1
CONTROL #1

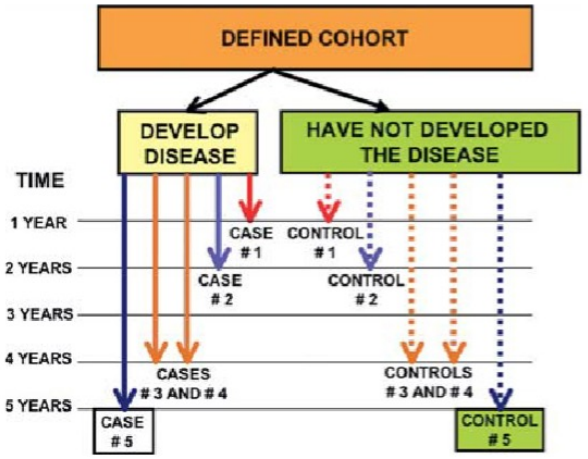

DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE 1
CONTROL 1
CASE #2


DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE 1
CONTROL 1
CASE #2
CONTROL #2


DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE #1
CONTROL #1
CASE #2
CONTROL #2
CASES #3 AND #4


DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE #1
CONTROL #1
CASE #2
CONTROL #2
CASES #3 AND #4
CONTROLS #3 AND #4


DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE #1
CONTROL #1
CASE #2
CONTROL #2
CASES #3 AND #4
CONTROLS #3 AND #4
CASE #5

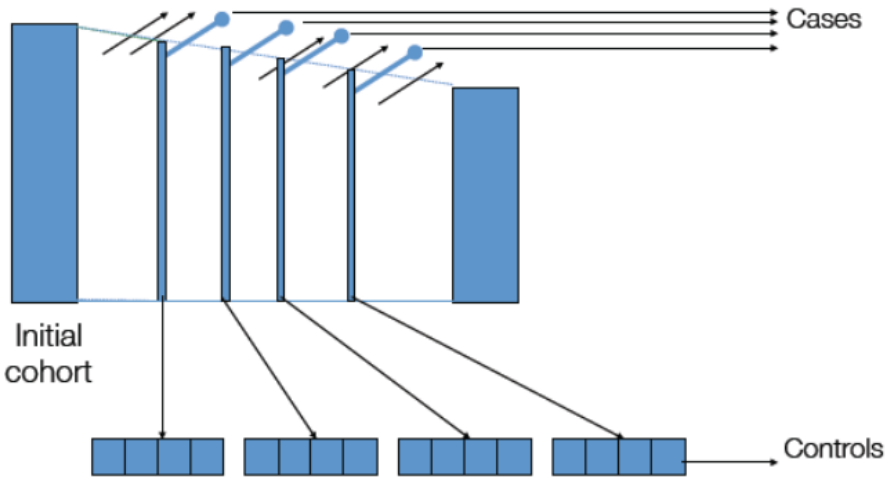

DEFINED COHORT
DEVELOP DISEASE
HAVE NOT DEVELOPED THE DISEASE
TIME
1 YEAR
2 YEARS
3 YEARS
4 YEARS
5 YEARS
CASE #1
CONTROL #1
CASE #2
CONTROL #2
CASES #3 AND #4
CONTROLS #3 AND #4
CASE #5
CONTROL #5

Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclo-oxygenase 2 selective and non-selective non-steroidal anti-inflammatory drugs: nested case-control study

David J Graham, David Campen, Rita Hui, Michele Spence, Craig Cheetham, Gerald Levy, Stanford Shoor, Wayne A Ray

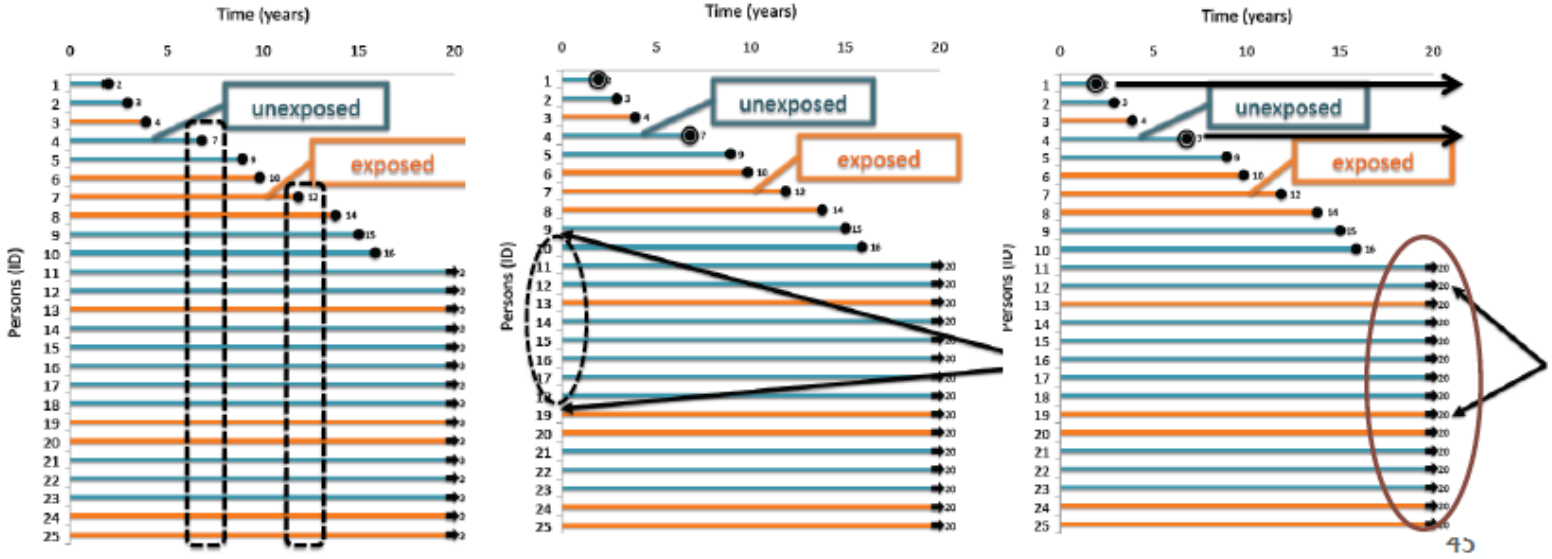
Identified a cohort of individuals age 18–84 years [~ 6 million] who filled at least one prescription for a COX2 selective (celecoxib or rofecoxib) or non-selective (all other) NSAID (1999-2000). Cases were individuals who developed serious coronary heart disease (acute MI or sudden cardiac death).

Randomly selected four controls from individuals under observation in the study cohort on the date of the case event (index date)…A given cohort member selected as a control for a case on one date could become a control for another case occurring on a later index date, as long as he or she remained in the study cohort and was therefore also at risk of becoming a case. Thus, a control could subsequently become a case.



Review: Case Control Designs

	1. Density case-control	2. Case-cohort	3. Cumulative case-control
Sampling from:	Non-cases at time of diagnosis ("risk sets")	Entire cohort at baseline (non-cases at baseline)	Non-cases at the end of follow-up
OR estimates:	Incidence Rate Ratio	Risk Ratio	Odds Ratio



Practical issues concerning selection of cases and controls

Defining Cases

- Ideally, cases in a case-control study will comprise all (or a representative sample) of members of a defined population who the outcome of interest during a specific period of time
- Eligibility criteria for cases must be carefully specified, just like in cohort studies and RCTs
- Incident cases vs. prevalent cases
  - Which should be used to study disease etiology?

Principles of Control Selection

Two important rules for control selection:

1. Controls must be selected independently of exposure

2. Controls must be selected from the source population that gave rise to the cases
   - Ideal= random sampling from source population
   - This is often where case-control studies run into difficulty

Types of study base

Can select controls from either a primary or secondary study base

1. Primary study base

- Cases are a representative sample of all cases in a defined population and controls are sampled directly from this source population
- The cases are subjects within the base who develop disease

2. Secondary study base

- Cases are selected before the study base is identified
- The study base then is defined as the source of the cases; controls are people who would have been recognized as cases if they had developed disease

Examples: primary or secondary base?

Imagine a study of Hep-C co-infection among men in the MACS study. Investigators recruit all men who develop Hep-C from 1990-1995.

Imagine a study of brain tumors at Roswell Park Hospital. The investigators recruit all incident cases of brain tumor during 2020.

Comparing Control Selection

- It is easier to sample controls from a primary study base
  - Well-defined
  - Cohort, population registry
- Control sampling is very difficult from a secondary study base because it is difficult to identify who is or isn't a member of the study base
- Risk of bias is much greater with a secondary study base than with a primary study base

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.

- Population controls
- Community/ neighbourhood controls
- Spouse/relative/friend controls
- Hospital controls

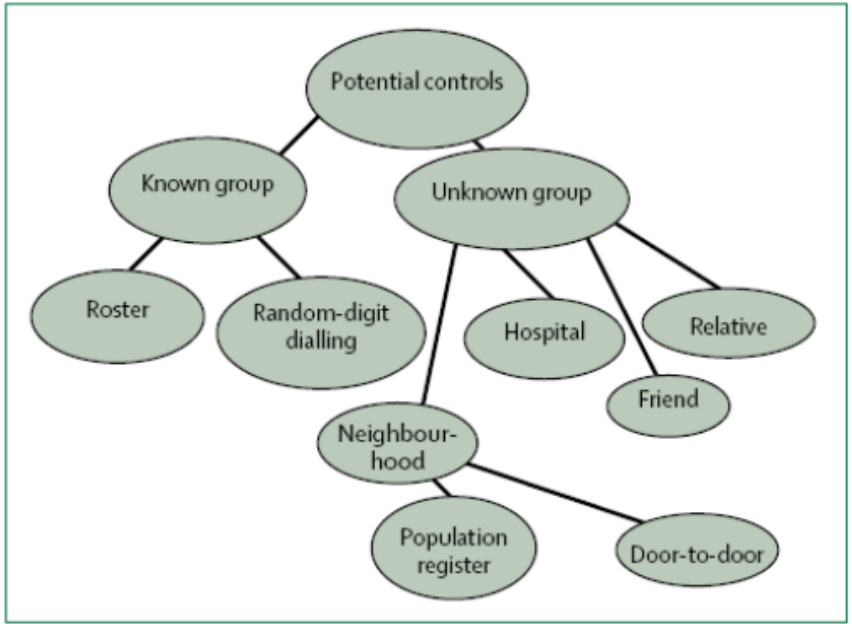


Figure 2: Choosing controls with known and unknown group of study participants

Population Controls

- In a population-based study, the source population is a geographically defined area (e.g., city, province etc.)
- When a population roster is available, selection of population controls is simplest
  - Census lists, drivers license records
- If no roster is available:
  - Random digit dialing
  - Neighborhood controls

Danish population-based registers for public health and health-related welfare research: Introduction to the supplement

LAU CASPAR THYGESEN & ANNETTE KJÆR ERSBØLL

European Centre for Register-Based Health-Related Population Research – Public Health, Major Diseases and Welfare (ECREPH), National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark

Denmark and other Nordic countries have exceptional opportunities to perform register-based research, because of the unique personal identification number available to all persons with permanent residence [1]. This number makes it possible to link information at the individual level from several registers for investigation of various research questions. The unique personal identification number was introduced in Denmark in 1968, which enables follow-up of individuals for decades. This supple-

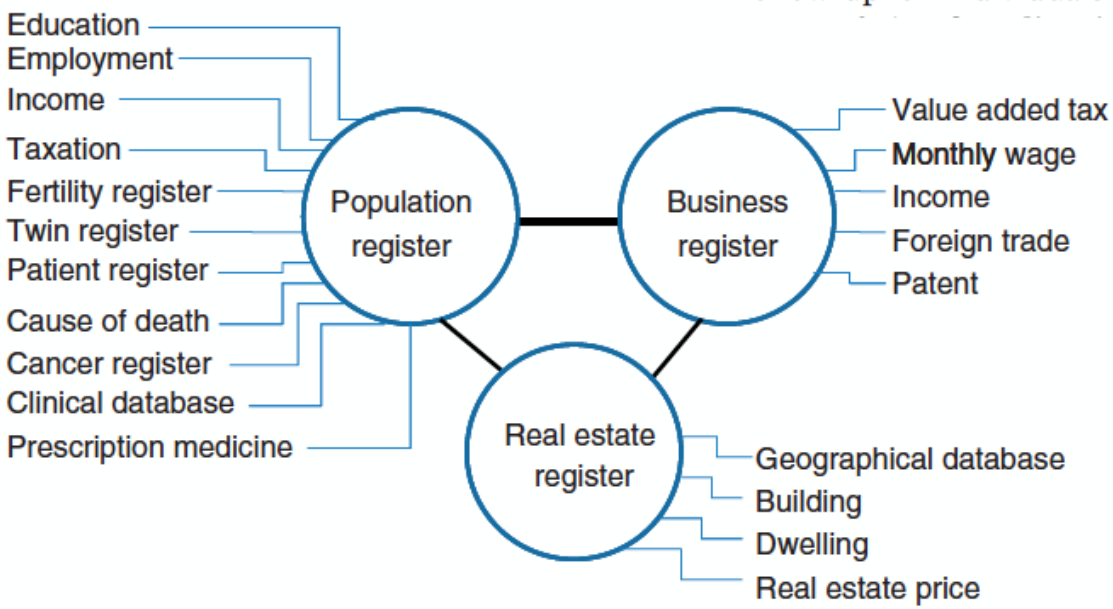
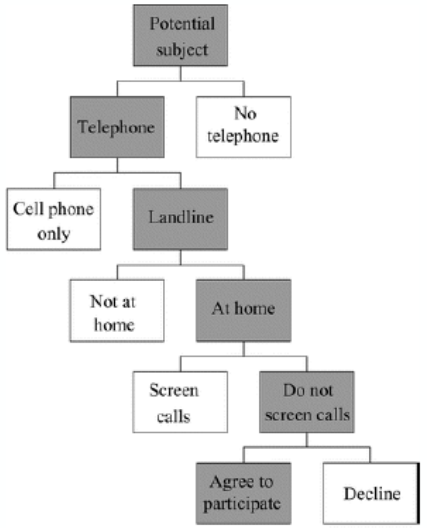


Figure 1. Schematic presentation of the base registers in Denmark. Linkage between all registers through three base registers. The keys are the personal identification number (CPR-number), the business identification number (SE/CVR-number) and the housing identification number (BBR-number). Inspired by Wallgren and Wallgren (2007) [4].

Random Digit Dialing

- Controls sampled in this way are mostly representative of the population
- However:
  - Not everyone has telephones
  - People at home during the day may not represent the pop'n
  - Caller ID
  - Low response
  - Cell phones?



Kempf AM, Remington PL. 2007. Annu. Rev. Public Health 28:113–26

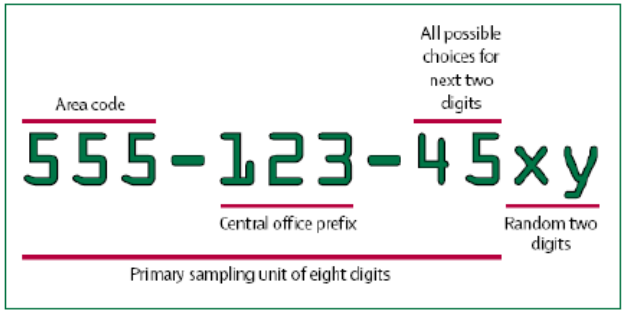


Figure 3: **Random-digit dialling for controls**[17]
Primary sampling unit included eight-digit numbers: all known area codes and three-digit central-office prefixes in the county, plus all combinations of next two digits. For all these eight-digit numbers randomly chosen, a computer generated the two final digits, creating a ten-digit number to be called.

Neighborhood Controls

- Sampling residences in a systematic way
- Drawing controls from the same neighborhood as the cases, certain confounding variables are accounted for (e.g., SES, climate)
- Sampling houses rather than individuals
- Non-response can be very high (like in RDD)
- Multi-unit dwellings

Friend/Relative Controls

- May be more willing to participate
- Could control for SES, ethnicity, genetics (?)
- However, friend/relative controls tend to have an exposure distribution that is more similar to cases than that of the source population
- Selection of controls may not be independent of exposure

Spouse Controls

OPEN ACCESS Freely available online

PLOS MEDICINE

Travel-Related Venous Thrombosis: Results from a Large Population-Based Case Control Study (MEGA Study)

Suzanne C. Cannegieter[1], Carine J. M. Doggen[1], Hans C. van Houwelingen[2], Frits R. Rosendaal[1,3*]

1 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands, 2 Department of Medical Statistics, Leiden University Medical Center, Leiden, Netherlands, 3 Thrombosis and Haemostasis Research Center, Department of Haematology, Leiden University Medical Center, Leiden, Netherlands

Funding: This research was supported by the Netherlands Heart Foundation (NHS 98.113) and the Dutch Cancer Foundation (RUL 99/1992). SCC is employed on the WRIGHT Project, which is being carried out under the auspices of the World Health Organization and funded jointly by the United Kingdom government and the European Commission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Academic Editor: Eduardo Franco, McGill University, Canada

Citation: Cannegieter SC, Doggen CJM, van Houwelingen HC, Rosendaal FR (2006) Travel-related venous thrombosis: Results from a large population-based case control study (MEGA study). PLoS Med 3(8): e307. DOI: 10.1371/journal.pmed.0030307

Received: March 21, 2005
Accepted: May 9, 2006
Published: August 22, 2006

ABSTRACT

Background

Recent studies have indicated an increased risk of venous thrombosis after air travel. Nevertheless, questions on the magnitude of risk, the underlying mechanism, and modifying factors remain unanswered.

Methods and Findings

We studied the effect of various modes and duration of travel on the risk of venous thrombosis in a large ongoing case-control study on risk factors for venous thrombosis in an unselected population (MEGA study). We also assessed the combined effect of travel and prothrombotic mutations, body mass index, height, and oral contraceptive use.

Since March 1999, consecutive patients younger than 70 y with a first venous thrombosis have been invited to participate in the study, with their partners serving as matched control individuals. Information has been collected on acquired and genetic risk factors for venous thrombosis. Of 1,906 patients, 233 had traveled for more than 4 h in the 8 wk preceding the event. Traveling in general was found to increase the risk of venous thrombosis 2-fold (odds ratio [OR] 2.1; 95% confidence interval [CI] 1.5–3.0). The risk of flying was similar to the risks of traveling by car, bus, or train. The risk was highest in the first week after traveling. Travel by car, bus, or train led to a high relative risk of thrombosis in individuals with factor V Leiden (OR 8.1; 95% CI 2.7–24.7), in those who had a body mass index of more than 30 kg/m^2 (OR 9.9; 95% CI 3.6–27.6), in those who were more than 1.90 m tall (OR 4.7; 95% CI 1.4–15.4), and in those who used oral contraceptives (estimated OR > 20). For air travel these synergistic findings were more apparent, while people shorter than 1.60 m had an increased risk of thrombosis after air travel (OR 4.9; 95% CI 0.9–25.6) as well.

Sibling Controls

6 Freeman J, McGowan Jr JE. Risk factors for nosocomial infection. *J Infect Dis* 1978;**138**:811–9
7 Kollef MH. Time to get serious about infection prevention in the ICU. *Chest* 2006;**130**:1293–6

Does breast feeding provide protection against acute appendicitis? A case-control study

João Guilherme Bezerra Alves MD PhD
José Natal Figueiroa BSc MSc
Isabela Barros

Instituto Materno Infantil Prof. Fernando Figueira (IMIP), Escola Pernambucana de Medicina (Faculdade Boa Viagem – FBV/IMIP), Recife, Brazil

Correspondence to: João Guilherme Bezerra Alves, Instituto Materno Infantil Prof. Fernando Figueira (IMIP), Rua dos Coelhos, 300, Boa Vista, Recife, Pernambuco, Brazil, CEP: 50.070-550
Email: joaoguilherme@imip.org.br

TROPICAL DOCTOR 2008; 38: 235–236
DOI: 10.1258/td.2008.070404

Methods and patients

The study population consisted of 243 children and adolescents who underwent surgery for suspected appendicitis at our hospitals – Instituto Materno Infantil Professor Fernando Figueira (IMPI) and Hospital da Restauracao (HR) – in Recife, northeast Brazil between 1 August 2006 and 30 March 2007. The 200 of these who did have histologically confirmed appendicitis were recruited to the study, matched by 200 familial controls (i.e. a sibling of the same gender and age within three years without a history of appendicitis). All the mothers were interviewed during the hospital stay. The study was approved by the ethics committee at the IMIP. We asked mothers how and for how long were their children – those with and those without appendicitis – fed milk during the first year of life and if they were breast feed only, given a mixture of breast and bottle or bottle feed only.

The sample size was based on the assumption that a 15% difference in the prevalence of breast feeding between the two groups would be clinically significant. The SPSS 12.0 for Windows (SPSS, Inc, Chicago, IL, USA) was used for the analysis of data. Quantitative data were expressed as means $\pm$ standard deviation (SD). Differences in continuous variables were analysed by the Mann-Whitney *U*-test or Student's *t*-test. Differences in categorical variables were assessed with the Fisher's exact test and the chi-squared test with Yate's correction: a P value <0.05 was considered statistically significant.

Hospital Controls

- Control group selected from patients treated at the same hospital as the cases
- Easily accessible population
  - Assumption is that patients treated for another disease would have also been treated for the disease under study at the same hospital
- Not a random sample of the source population, possible that controls are not selected independently of exposure

Coffee and Pancreatic Cancer

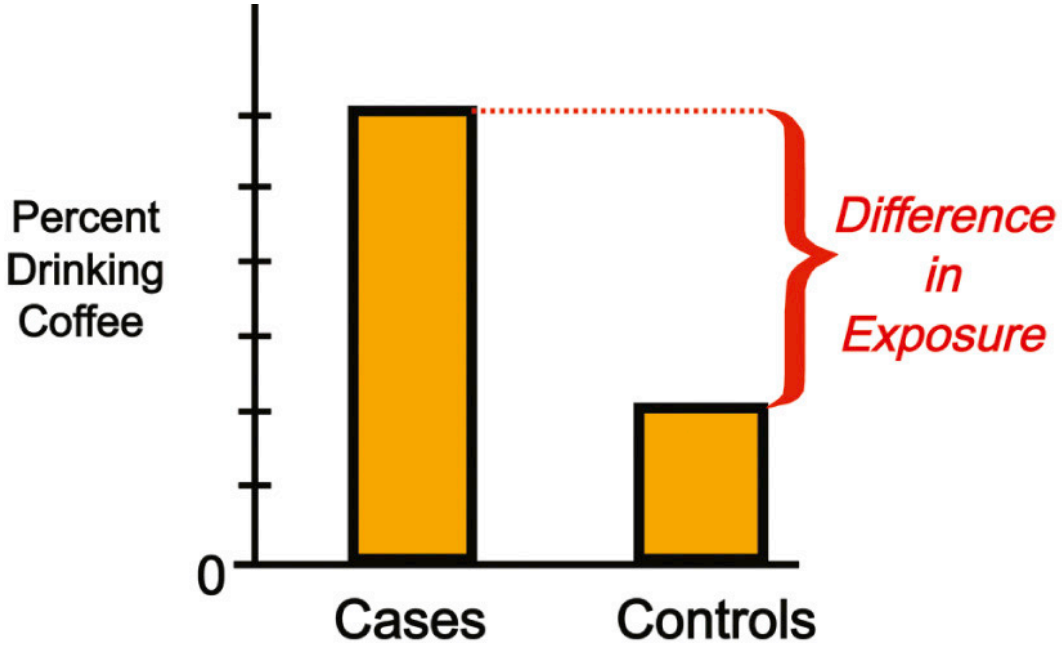
COFFEE AND CANCER OF THE PANCREAS

BRIAN MACMAHON, M.D., STELLA YEN, M.D., DIMITRIOS TRICHOPOULOS, M.D., KENNETH WARREN, M.D., AND GEORGE NARDI, M.D.

Abstract We questioned 369 patients with histologically proved cancer of the pancreas and 644 control patients about their use of tobacco, alcohol, tea, and coffee. There was a weak positive association between pancreatic cancer and cigarette smoking, but we found no association with use of cigars, pipe tobacco, alcoholic beverages, or tea. A strong association between coffee consumption and pancreatic cancer was evident in both sexes. The association was not affected by controlling for cigarette use. For the sexes combined, there was a significant dose-response relation ($P \sim 0.001$); after adjustment for cigarette smoking, the relative risk associated with drinking up to two cups of coffee per day was 1.8 (95 per cent confidence limits, 1.0 to 3.0), and that with three or more cups per day was 2.7 (1.6 to 4.7). This association should be evaluated with other data; if it reflects a causal relation between coffee drinking and pancreatic cancer, coffee use might account for a substantial proportion of the cases of this disease in the United States. (N Engl J Med. 1981; 304:630-3.)

- Controls selected from a group of patients hospitalized by the same physicians who had diagnosed and hospitalized the cases' disease
- Investigators attempted to make the selection process similar for cases and controls
- However, one of the first things told to any GI patient is "reduce coffee intake"
- Controls had an unusually low prevalence of exposure (coffee intake)
- Results of McMahon study not replicated when population based controls were used

Do patients with pancreatic cancer drink more coffee than people without pancreatic cancer in the same population?



- In this study, the level of coffee drinking in cases was greater than the level of coffee drinking in controls
  - Controls' levels of coffee drinking ≠ of the level of coffee drinking in the population
    - Coffee drinking may be abnormally low.
    - Observed difference in coffee drinking between pancreatic cancer cases and controls could be due to: cases drinking more coffee than expected **or** controls drinking less coffee than expected

Composing a hospital control series

- Exclude from the control series any hospitalizations for any conditions related to the exposure
- Example: case control study of NSAIDs and colorectal cancer

Panel 2: **Introduction of bias through poor choice of controls**

Cases	Control selection	Non-representativeness
Colorectal cancer patients admitted to hospital	Patients admitted to hospital with arthritis	Controls probably have high degrees of exposure to NSAIDs
Colorectal cancer patients admitted to hospital	Patients admitted to hospital with peptic ulcers	Controls probably have low degrees of exposure to NSAIDs

NSAIDs=non-steroidal anti-inflammatory drugs.

- Prior history of disease should not exclude control subjects unless the same restriction put on cases
- Patients with any disease that is not easily distinguishable from the study disease should be excluded

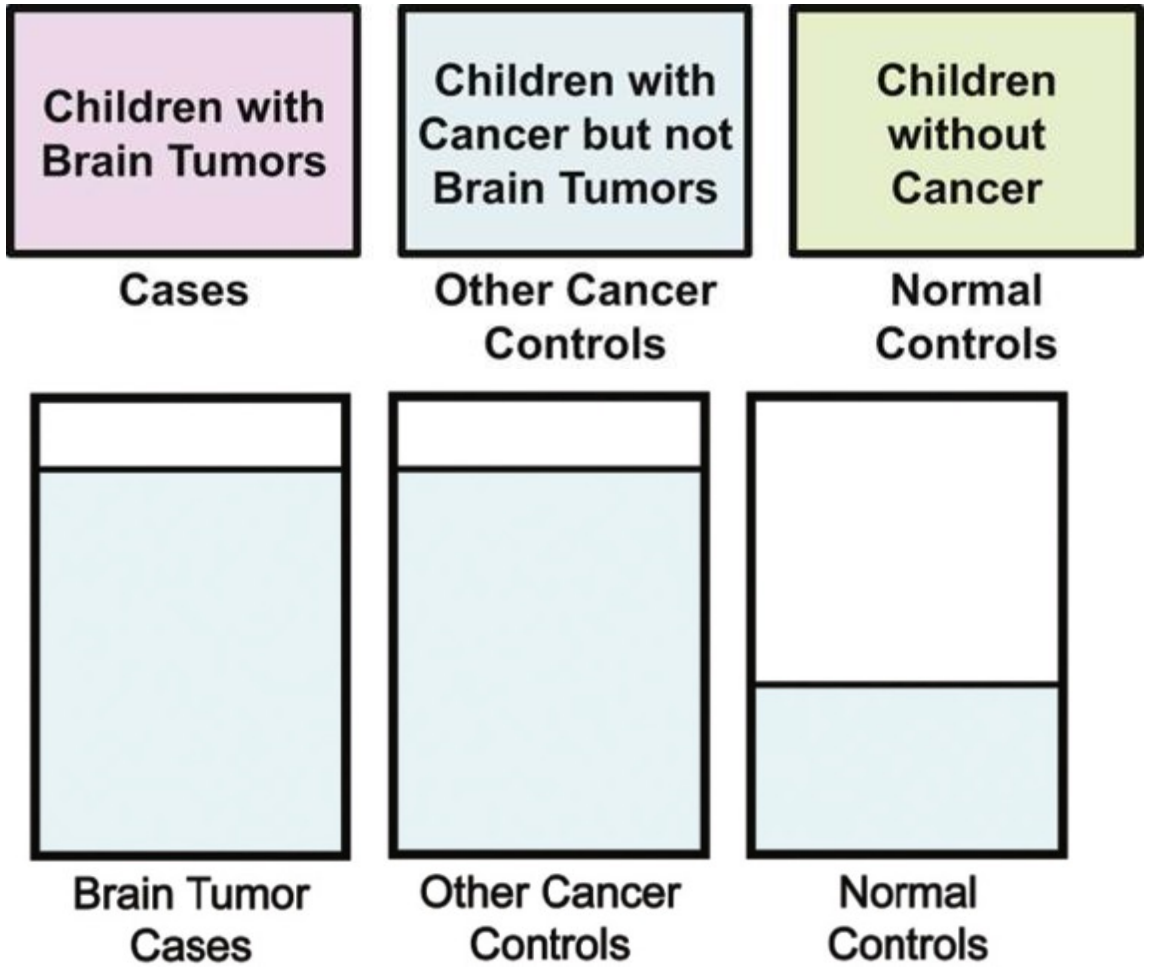
Multiple Control Groups

- Some researchers have suggested using multiple control groups to examine alternate hypotheses and potential sources of bias
- Multiple controls of same type (e.g., 1:2 or 1:4)
- Multiple controls of different type (e.g., hospital and neighborhood)
  - Reassuring when results are concordant regardless of which control group is used
  - When the results are discordant, it puts the investigator in a position to choose which is ‘most correct’

What is the ideal control group?

- Some researchers suggest using both
- Infer that the true effect estimate is somewhere between the two estimates

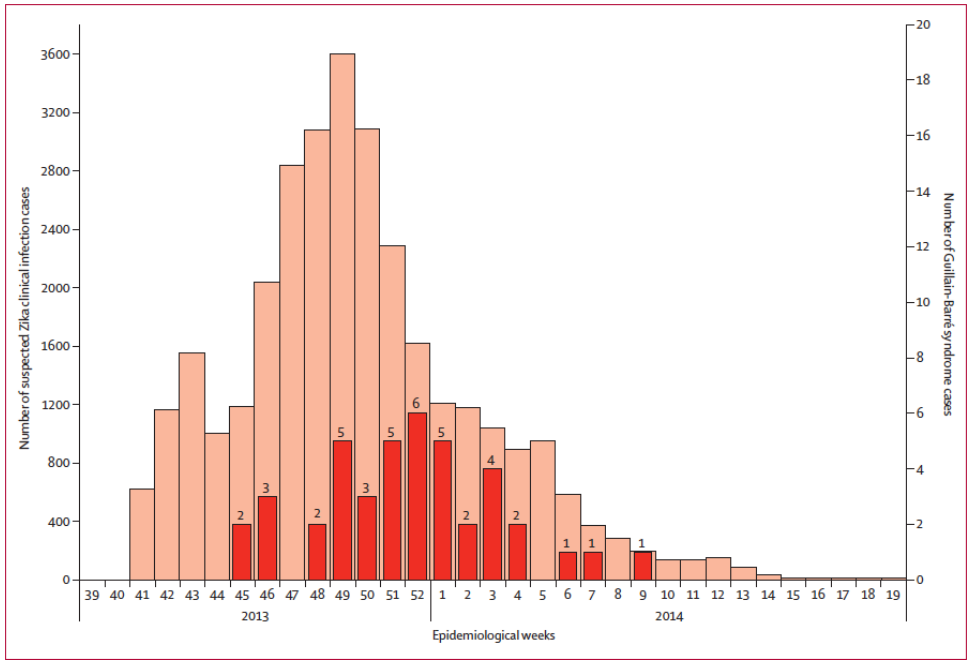
Did mothers of children with brain tumors have more prenatal radiation exposure than control mothers?



Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study

Van-Mai Cao-Lormeau, Alexandre Blake*, Sandrine Mons, Stéphane Lastère, Claudine Roche, Jessica Vanhomwegen, Timothée Dub, Laure Baudouin, Anita Teissier, Philippe Larre, Anne-Laure Vial, Christophe Decam, Valérie Choumet, Susan K Halstead, Hugh J Willison, Lucile Musset, Jean-Claude Manuguerra, Philippe Despres, Emmanuel Fournier, Henri-Pierre Mallet, Didier Musso, Arnaud Fontanet*, Jean Neil*, Frédéric Ghawché**

Background Between October, 2013, and April, 2014, French Polynesia experienced the largest Zika virus outbreak ever described at that time. During the same period, an increase in Guillain-Barré syndrome was reported, suggesting a possible association between Zika virus and Guillain-Barré syndrome. We aimed to assess the role of Zika virus and dengue virus infection in developing Guillain-Barré syndrome.



Methods

Study design and participants

In this case-control study, cases were patients with Guillain-Barré syndrome who were diagnosed at the Centre Hospitalier de Polynésie Française (CHPF) in Papeete, Tahiti, French Polynesia, during the outbreak period. As routine, all patients with suspicion of Guillain-Barré syndrome in French Polynesia are referred to the CHPF for diagnosis confirmation. All patients included in this study were diagnosed as developing a Guillain-Barré syndrome by neurologists or staff in intensive care

Case Series

(control group 1; n=98) was recruited among patients in hospital or consulting for, non-febrile illness at the CHPF. Patients from the control group 1 were matched for age (within a 10-year margin), sex, and island of residence with patients in the Guillain-Barré syndrome

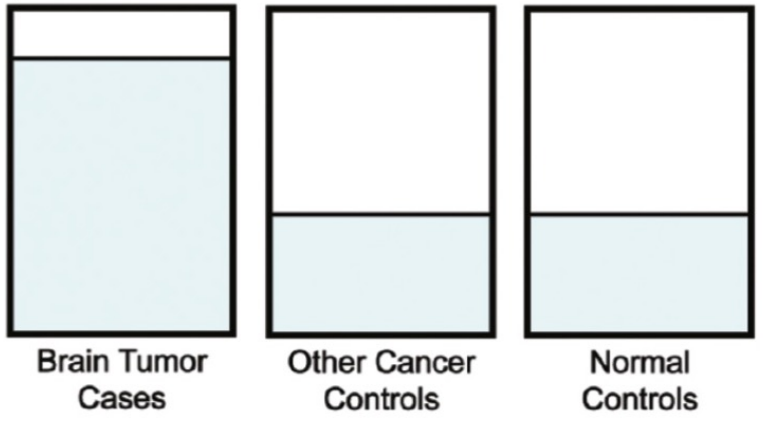
Control group 1

To investigate a possible role of past dengue infection in developing Guillain-Barré syndrome in a Zika virus infected patient, a second control group (control group 2; n=70) was recruited among age-matched (within a 10-year margin) patients with RT-PCR-confirmed Zika virus infection, but who did not develop any neurological complication.

Control group 2

Bias in Case Control Studies

- Case control studies are particularly susceptible to bias
  - Will discuss in further detail in the Bias lecture
- Poor recall vs. recall bias
  - Poor memory can happen with both cases and controls
  - Cases often have different recall than controls ("rumination bias") e.g., congenital malformations
  - Sometimes cases are probed/investigated differently

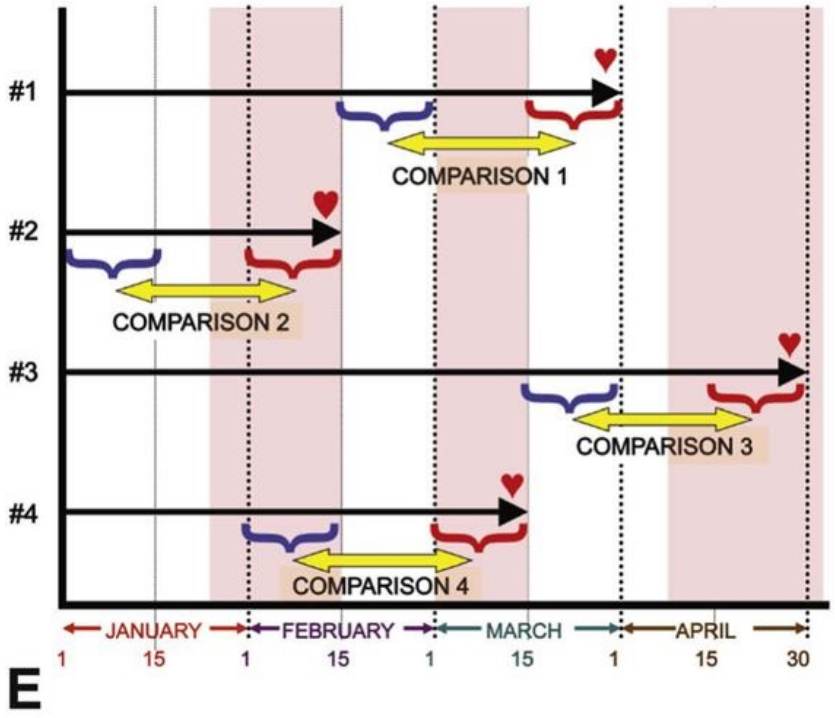


Matching

- A frequent concern in case control studies is that cases and controls may differ in characteristics aside from exposure
- Selecting cases and controls so they are similar on particular factors of interest (e.g., age)
- Frequency matching
  - Same proportion (e.g., 10%) of older adults in cases and controls
- Individual matching
  - For each older adult case, an older adult control is selected

Case-Crossover Study

- Useful to study etiology of **acute events** in situations where the exposure is transient and its effect occurs over a short time
  - Subjects must cross back and forth between periods of risk
- Each case serves as his or her own control(s)
- Not concerned about other differences between the characteristics of the cases and those of a separate group of controls.
- This design also eliminates the additional cost that would be associated with identifying and interviewing a separate control population.



In this type of study, a case is identified (for example, a person who had a heart attack) and the level of exposure is ascertained for a short time period preceding the event (the at-risk period).

This level is compared with the level of exposure in a control time period that is more remote from the event.

Abstract

Background Because of a belief that the use of cellular telephones while driving may cause collisions, several countries have restricted their use in motor vehicles, and others are considering such regulations. We used an epidemiologic method, the case–crossover design, to study whether using a cellular telephone while driving increases the risk of a motor vehicle collision.

Methods We studied 699 drivers who had cellular telephones and who were involved in motor vehicle collisions resulting in substantial property damage but no personal injury. Each person's cellular-telephone calls on the day of the collision and during the previous week were analyzed through the use of detailed billing records.

Results A total of 26,798 cellular-telephone calls were made during the 14-month study period. The risk of a collision when using a cellular telephone was four times higher than the risk when a cellular telephone was not being used (relative risk, 4.3; 95 percent confidence interval, 3.0 to 6.5). The relative risk was similar for drivers who differed in personal characteristics such as age and driving experience; calls close to the time of the collision were particularly hazardous (relative risk, 4.8 for calls placed within 5 minutes of the collision, as compared with 1.3 for calls placed more than 15 minutes before the collision; $P<0.001$); and units that allowed the hands to be free (relative risk, 5.9) offered no safety advantage over hand-held units (relative risk, 3.9; P not significant). Thirty-nine percent of the drivers called emergency services after the collision, suggesting that having a cellular telephone may have had advantages in the aftermath of an event.

The New England Journal of Medicine

VOLUME 336 FEBRUARY 13, 1997 NUMBER 7

ASSOCIATION BETWEEN CELLULAR-TELEPHONE CALLS AND MOTOR VEHICLE COLLISIONS

DONALD A. REDELMEIER, M.D., AND ROBERT J. TIBSHIRANI, PH.D.

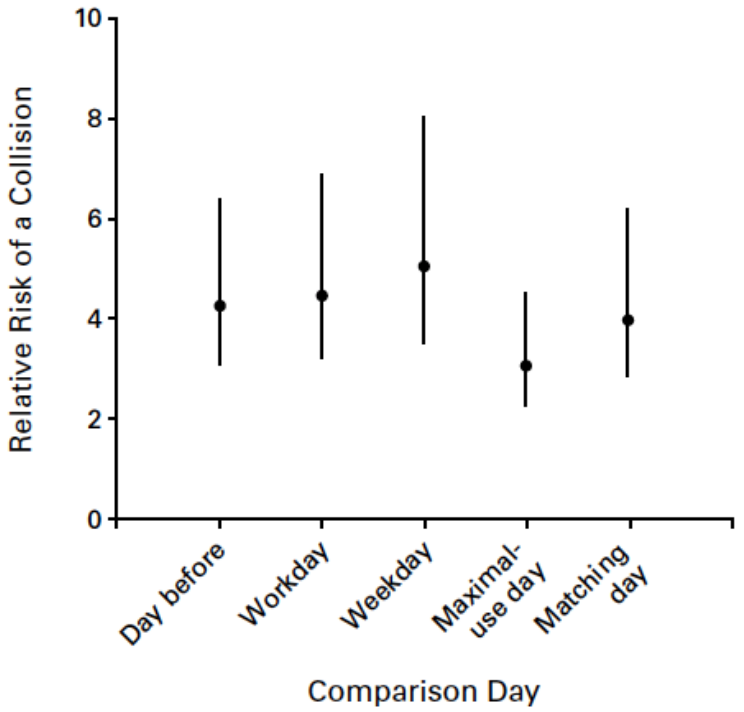


Figure 1. Relative Risk of a Collision for Different Control Periods.