

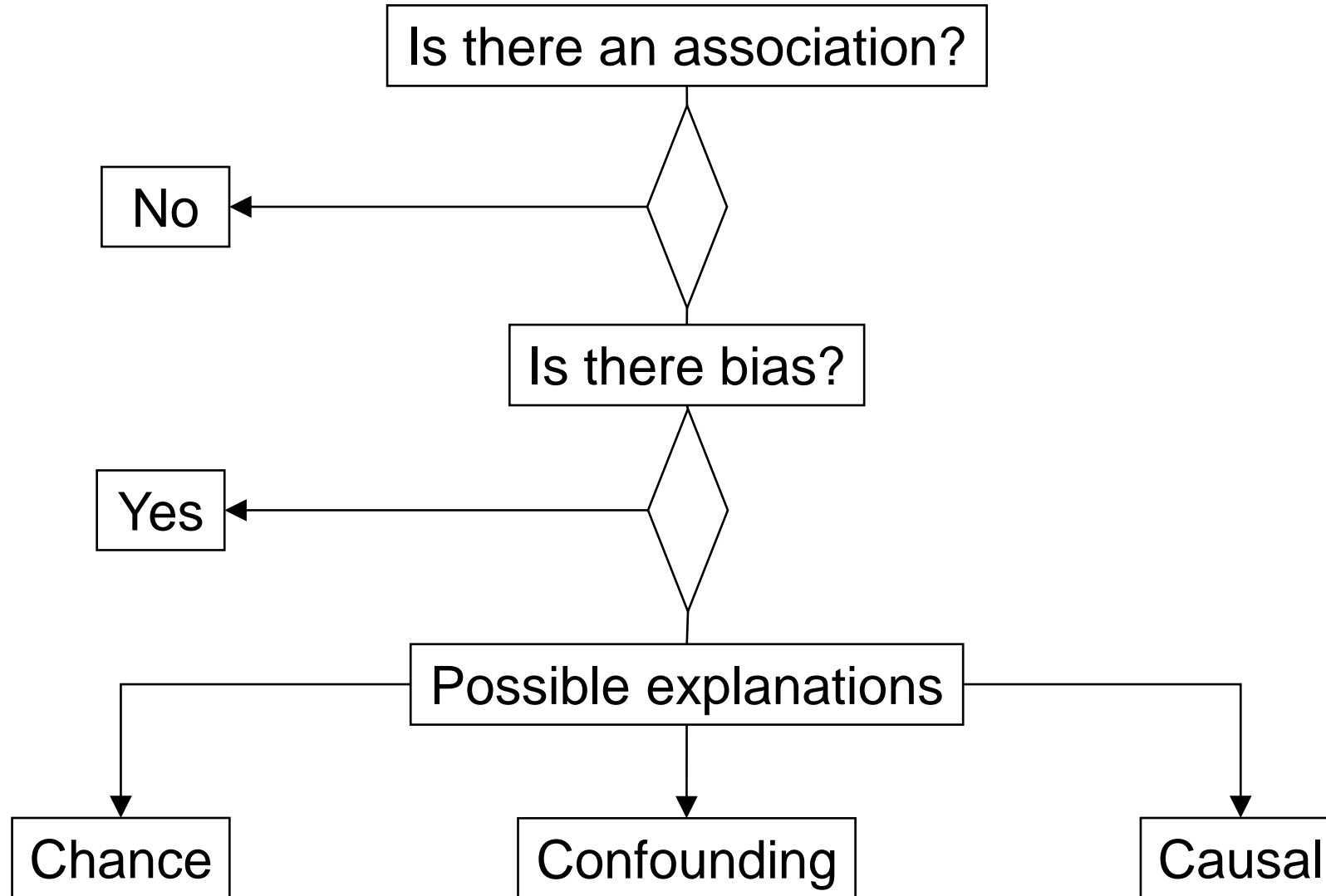
Overview of Study Designs

Kristen A. Stafford

Associate Professor

Department of Epidemiology and Public Health

Epidemiologic Reasoning



Epidemiologic Study Designs

Experimental Studies

- Randomized Controlled Trials
- Other Experimental Studies

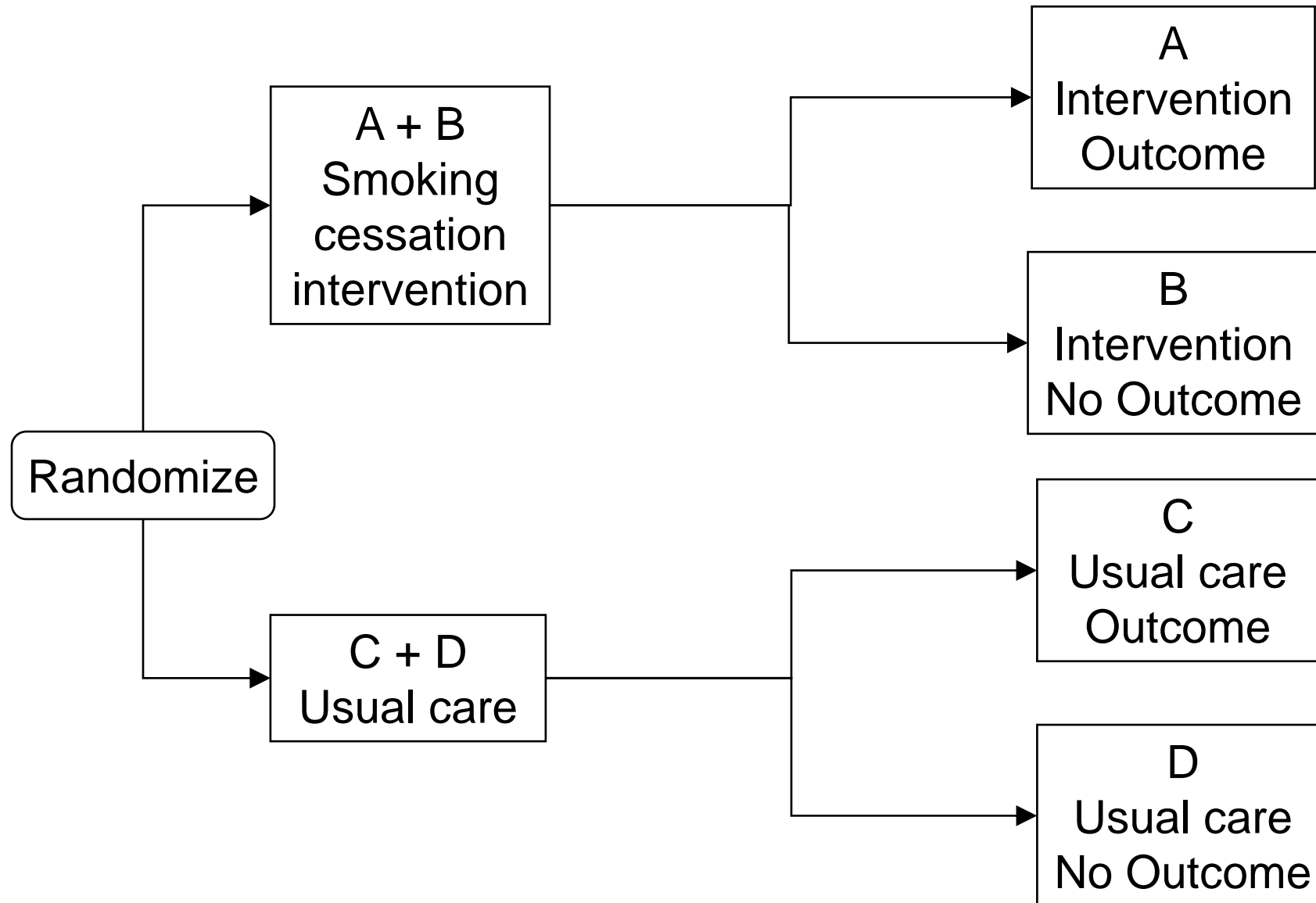
Observational Studies

- Cohort Studies
- Case-Control Studies
- Cross-Sectional Studies
- Ecologic Studies
- Case Series

Randomized Controlled Trials

- Treated and untreated subjects are followed over time to determine whether they experience the outcome (e.g., relapse, death, clinical improvement)
- Assignment to treatment or non-treatment is by randomization

Randomized Controlled Trials



Randomized Controlled Trials

		Outcome		
		Yes	No	
Intervention	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	

$$\text{Relative risk} = \frac{\text{Outcome risk in intervention group}}{\text{Outcome risk in usual care group}} = \frac{a/(a+b)}{c/(c+d)}$$

RCT Questions

- How did they randomize patients?
- Could allocation be predicted?
- Are groups fairly balanced with respect to covariates (check Table 1)?
- Was there a lot of lost to follow up?
- Did they perform an intention to treat analysis?
- Were outcomes assessed in the same way across groups?
- Was the study appropriately powered
- Are the results generalizable to your patient population?

The screenshot shows a web browser window with the URL www.consort-statement.org. The navigation menu includes Home, CONSORT 2010, Extensions, Downloads, Examples, Resources, and About CONSORT. The main content area features a large banner with the CONSORT logo and the text "TRANSPARENT REPORTING of TRIALS". Below the banner, a welcome message states: "Welcome to the CONSORT Website. CONSORT stands for Consolidated Standards of Reporting Trials and encompasses various initiatives developed by the CONSORT Group to alleviate the problems arising from inadequate reporting of randomized controlled trials." A section titled "The CONSORT Statement" explains that the main product is the CONSORT Statement, which provides a standard way for authors to prepare reports of trial findings. On the right side, there is a "CONSORT 2010 Key Documents" section with links to the CONSORT 2010 Checklist, Flow Diagram, Statement, and Explanation and Elaboration Document. Below this is a "Tweets by @CONSORTing" section featuring a tweet from Toby Lasserson (@tlassers) discussing compliance with RCT reporting standards. The Windows taskbar at the bottom shows various application icons and the system clock indicating 11:22 AM on 7/10/2016.

<http://www.consort-statement.org/>



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	_____
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	_____
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	_____
	2b	Specific objectives or hypotheses	_____
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	_____
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	_____
Participants	4a	Eligibility criteria for participants	_____
	4b	Settings and locations where the data were collected	_____
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	_____
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	_____
	6b	Any changes to trial outcomes after the trial commenced, with reasons	_____
Sample size	7a	How sample size was determined	_____
	7b	When applicable, explanation of any interim analyses and stopping guidelines	_____
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	_____
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	_____
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	_____
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	_____
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	_____

Examples of RCTs

Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33)

UK Prospective Diabetes Study (UKPDS) Group*

Published: September 12, 1998 • DOI: [https://doi.org/10.1016/S0140-6736\(98\)07019-6](https://doi.org/10.1016/S0140-6736(98)07019-6)

ORIGINAL ARTICLE

Prevention of HIV-1 Infection with Early Antiretroviral Therapy

Myron S. Cohen, M.D., Ying Q. Chen, Ph.D., Marybeth McCauley, M.P.H., Theresa Gamble, Ph.D., Mina C. Hosseinipour, M.D., Nagalingeswaran Kumarasamy, M.B., B.S., James G. Hakim, M.D., Johnstone Kumwenda, F.R.C.P., Beatriz Grinsztejn, M.D., Jose H.S. Pilotto, M.D., Sheela V. Godbole, M.D., Sanjay Mehendale, M.D., *et al.*, for the HPTN 052 Study Team*

The New England Journal of Medicine

©Copyright, 1994, by the Massachusetts Medical Society

Volume 331

NOVEMBER 3, 1994

Number 18

REDUCTION OF MATERNAL–INFANT TRANSMISSION OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 WITH ZIDOVUDINE TREATMENT

EDWARD M. CONNOR, M.D., RHODA S. SPERLING, M.D., RICHARD GELBER, PH.D., PAVEL KISELEV, PH.D.,

Quasi-Experimental Studies

- Sometimes called pre-post/before-after intervention
- Non-randomized intervention studies
- Used to evaluate the effectiveness of specific interventions
- Often used for Quality Improvement initiatives
- Used increasingly in medical fields
- Social sciences full of examples

Key Questions

- Why did the authors choose a quasi-experimental design?
- Is the temporal sequence clear?
- Did they clearly identify which quasi design they used?
- Are there systematic differences in respondent characteristics that could cause the observed effect?
- Are there a large number of concurrent activities?
- Is this maturation (naturally occurring effect)?
- Could the effect be due to regression to the mean?
- How much attrition is there?
- Is there a practice effect?
- Did measurement change over time?
- How did they analyze their data?

A. QUASI-EXPERIMENTAL DESIGNS THAT DO NOT USE CONTROL GROUPS

1. The 1-group pretest-posttest design:

O1 X O2

2. The 1-group pretest-posttest design that uses a double pretest:

O1 O2 X O3

3. The 1-group pretest-posttest design that uses a nonequivalent dependent variable:

(O1a, O1b) X (O2a, O2b)

4. The removed-treatment design:

O1 X O2 O3 removeX O4

5. The repeated-treatment design:

O1 X O2 removeX O3 X O4

B. QUASI-EXPERIMENTAL DESIGNS THAT USE CONTROL GROUPS

0. The posttest-only design that uses nonequivalent groups:

X O1

O2

1. The untreated-control group design that uses dependent pretest and posttest samples:

O1a X O2a

O1b O2b

2. The untreated-control group design that uses dependent pretest and posttest samples and a double pretest:

O1a O2a X O3a

O1b O2b O3b

3. The untreated-control group design that uses dependent pretest and posttest samples and switching replications:

O1a X O2a O3a

O1b O2b X O3b

Examples of Quasi-Experimental Studies

[J Am Med Inform Assoc.](#) 2001 Mar-Apr; 8(2): 111–116.

doi: [10.1136/jamia.2001.0080111](https://doi.org/10.1136/jamia.2001.0080111)

PMCID: PMC134550

PMID: [11230379](https://pubmed.ncbi.nlm.nih.gov/11230379/)

Educational Instruction on a Hospital Information System for Medical Students During Their Surgical Rotations

[Robert Patterson](#), MD, MSc and [Peter Harasym](#), PhD

▸ [Author information](#) ▸ [Article notes](#) ▸ [Copyright and License information](#) [Disclaimer](#)

This article has been [cited by](#) other articles in PMC.

▸ [Int J Med Inform.](#) 2001 Oct;63(3):169-78. doi: [10.1016/s1386-5056\(01\)00177-0](https://doi.org/10.1016/s1386-5056(01)00177-0).

Implementation of clinical guidelines through an electronic medical record: physician usage, satisfaction and assessment

[V J Mikulich](#)¹, [Y C Liu](#), [J Steinfeldt](#), [D L Schriger](#)

Affiliations + expand

PMID: 11502431 DOI: [10.1016/s1386-5056\(01\)00177-0](https://doi.org/10.1016/s1386-5056(01)00177-0)

[J Am Med Inform Assoc.](#) 2003 Mar-Apr; 10(2): 177–187.

doi: [10.1197/jamia.M1175](https://doi.org/10.1197/jamia.M1175)

PMCID: PMC150371

PMID: [12595407](https://pubmed.ncbi.nlm.nih.gov/12595407/)

The Effect of Computer-generated Reminders on Charting Deficiencies in the ICU

[Thomas A. Oniki](#), PhD, [Terry P. Clemmer](#), MD, and [T. Allan Pryor](#), PhD

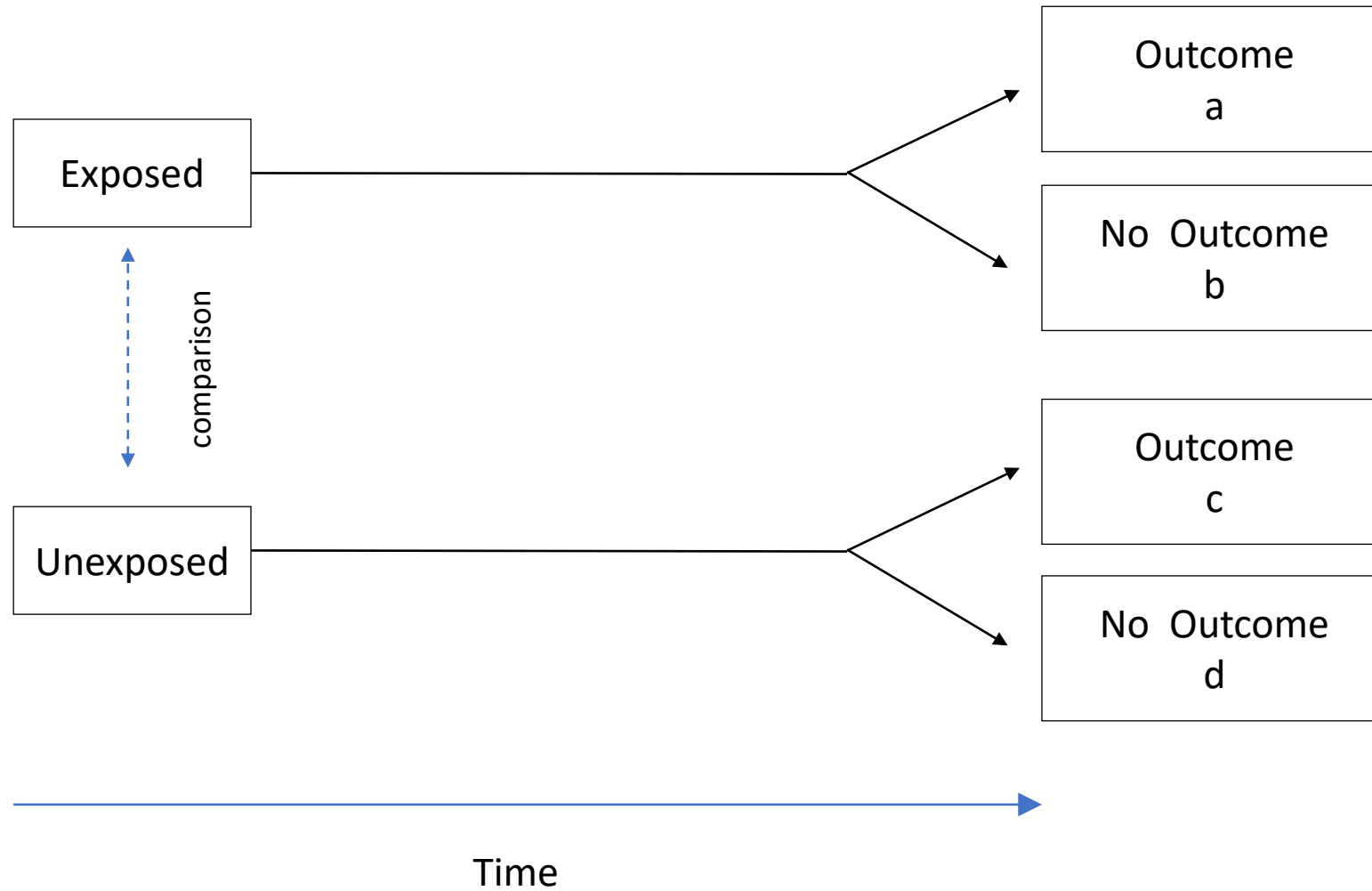
▸ [Author information](#) ▸ [Article notes](#) ▸ [Copyright and License information](#) [Disclaimer](#)

This article has been [cited by](#) other articles in PMC.

Cohort Studies

- Exposed and unexposed subjects without disease are followed over time to determine whether they experience the outcome
- Randomized controlled trials are a special case of the cohort study

Cohort Study



Prospective versus retrospective

- Usually deals with the time course over which data is collected
 - Prospective
 - People are recruited, found to be free of the outcome of interest, and followed over time (usually very long periods)
 - Nurses' Health Study (1976)
 - Framingham Heart Study (1948)
 - NA-ACCORD (2006)
 - WIHS (1993)
 - Retrospective
 - An existing source of health information (data) is used to retrospectively construct a cohort
 - The “time-course” between exposure and outcome is still prospective but the data is assembled and analyzed after events have happened
 - The retrospective is “looking back”

Cohort Studies

		Outcome		
		Yes	No	
Exposure	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	

$$\text{Relative risk} = \frac{\text{Risk of outcome among exposed}}{\text{Risk of outcome among unexposed}} = \frac{a/(a+b)}{c/(c+d)}$$

Cohort Questions

- How was exposure measured/defined?
- Were subjects at risk for development of outcome?
- Were outcomes assessed equally across exposure groups
- Does the study sample represent the source population (selection bias, internal validity)?
- Is loss to follow up/mortality informative (missing data, selection bias)?
- Could the exposure or outcome have been misclassified (information bias)?
- How was confounding assessed/controlled for?



STROBE Statement

Strengthening the reporting of observational studies in epidemiology

u^b

UNIVERSITÄT
DUISBURG
ESSEN

Home

Aims

News

Available checklists

Publications

Translations

Commentaries

Discussion forum

STROBE group

Endorsement

Contact

Links

Member login / logout

What is STROBE?

STROBE stands for an international, collaborative initiative of epidemiologists, methodologists, statisticians, researchers and journal editors involved in the conduct and dissemination of observational studies, with the common aim of **STrengthening the Reporting of OBservational studies in Epidemiology**.

The STROBE Statement is being endorsed by a growing number of biomedical journals. Click [here](#) for full list.

For STROBE-related entries in PubMed click [here](#).

What's new in the STROBE Initiative?

Observational Studies: Getting clear about transparency

01.09.2014

New guidelines for observational studies in PLOS Medicine [\[more\]](#)

[\[more\]](#)

New article of interest

01.07.2014

A Review of Published Analyses of Case-Cohort Studies and Recommendations for Future Reporting [\[more\]](#)

[\[more\]](#)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants (b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses

Examples of Cohort Studies



Nurses'
Health Study



Survival after the onset of congestive heart failure in Framingham Heart Study subjects.

K K Ho, K M Anderson, W B Kannel, W Grossman, and D Levy

Originally published 1 Jul 1993 | <https://doi.org/10.1161/01.CIR.88.1.107> | Circulation. 1993;88:107–115

Diet, smoking, social class, and body mass index in the Caerphilly Heart Disease Study

A M Fehily ✉, K M Phillips, J W Yarnell

The American Journal of Clinical Nutrition, Volume 40, Issue 4, October 1984, Pages 827–833, <https://doi.org/10.1093/ajcn/40.4.827>

Published: 01 October 1984 **Article history** ▼

ORIGINAL ARTICLE

HIV Prevention Efforts and Incidence of HIV in Uganda

M. Kate Grabowski, Ph.D., David M. Serwadda, M.B., Ch.B., M.P.H., Ronald H. Gray, M.D., Gertrude Nakigozi, M.B., Ch.B., Ph.D., Godfrey Kigozi, M.B., Ch.B., Ph.D., Joseph Kagaayi, M.B., Ch.B., Ph.D., Robert Ssekubugu, M.S.P.H., Fred Nalugoda, Ph.D., Justin Lessler, Ph.D., M.H.S., Thomas Lutalo, Ph.D., Ronald M. Galiwango, M.B., Ch.B., Sc.M., Fred Makumbi, Ph.D., *et al.*, for the Rakai Health Sciences Program*

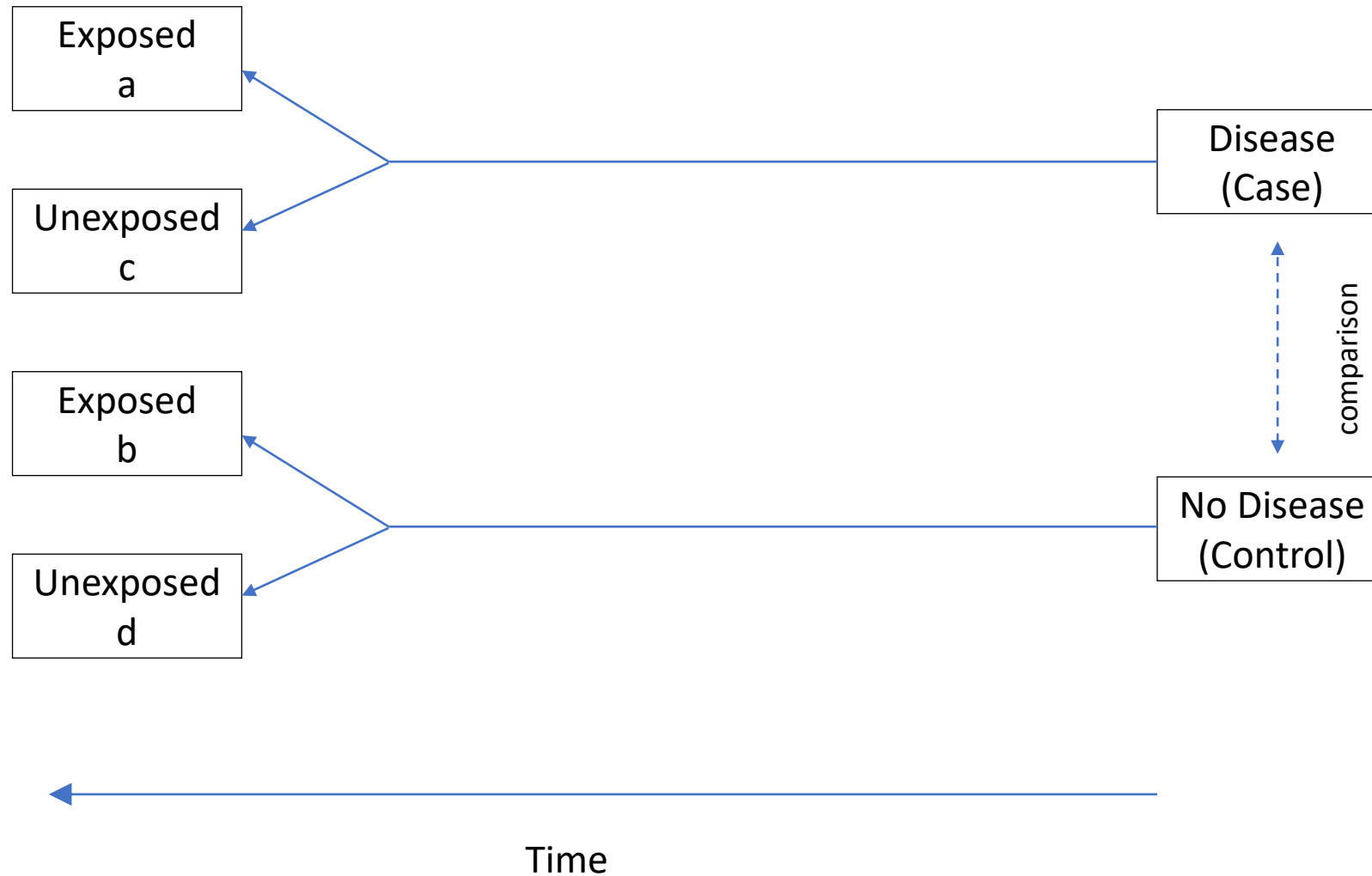
Case-Control Studies

- Compare exposure among persons with the disease (cases) to exposure among persons without the disease (controls)
- Most commonly used epidemiologic study design despite many potential biases
 - If not designed well
- If designed well, can be thought of as an efficient cohort study
 - Measures of association can approximate rate ratios or risk ratios

Case-Control Studies

- More efficient than the equivalent cohort study
- Makes it possible to study rare diseases
- Makes it possible to study diseases that take a long time to develop
- Used for outbreak investigations

Case-Control Study



Case-Control Studies

		Case	
		Yes	No
Exposed	Yes	a	b
	No	c	d
		a+c	b+d

$$\text{Odds ratio} = \frac{\text{Odds of exposure among cases}}{\text{Odds of exposure among controls}} = \frac{a/c}{b/d} = \frac{ad}{bc}$$

Questions for a Case-Control Study

- Was there a pre-specified hypothesis defining a relationship between an exposure and an outcome?
- Were the exposure and health outcome clearly and operationally defined?
- Was the control group appropriate?
- Was the measurement of exposure both in the cases and controls accurate and unbiased?
- Was the measurement of the outcome both in the cases and controls accurate and unbiased?
- Were the important confounding variables accounted for and controlled for in the statistical analysis?

Examples of Case-Control Studies

BRITISH MEDICAL JOURNAL

LONDON SATURDAY SEPTEMBER 30 1950

SMOKING AND CARCINOMA OF THE LUNG PRELIMINARY REPORT

BY

RICHARD DOLL, M.D., M.R.C.P.

Member of the Statistical Research Unit of the Medical Research Council

AND

A. BRADFORD HILL, Ph.D., D.Sc.

Professor of Medical Statistics, London School of Hygiene and Tropical Medicine; Honorary Director of the Statistical Research Unit of the Medical Research Council

Weekly

August 15, 1997 / 46(32);741-744

Persons using assistive technology might not be able to fully access information in this file. For assistance, please send e-mail to: mmwrq@cdc.gov. Type 508 Accommodation and the title of the report in the subject line of e-mail.

Outbreaks of Escherichia coli O157:H7 Infection Associated with Eating Alfalfa Sprouts -- Michigan and Virginia, June-July 1997

In June and July 1997, simultaneous outbreaks of Escherichia coli O157:H7 infection in Michigan and Virginia were independently associated with eating alfalfa sprouts grown from the same seed lot. The outbreak strains in Michigan and Virginia were indistinguishable by molecular subtyping methods. This report summarizes the preliminary findings of the outbreak investigations. Michigan

Cross-Sectional Studies

- Study in which the status of individuals with respect to one or more characteristics is assessed at one point in time

Cross-Sectional Studies

- May not be possible to determine whether exposure preceded disease
- No distinction between new cases and existing cases
- Not useful for the study of etiologic factors

Examples of Cross-Sectional Studies

The screenshot shows the CDC's Behavioral Risk Factor Surveillance System (BRFSS) website. At the top left is the CDC logo with the tagline "Centers for Disease Control and Prevention" and "CDC 24/7 Saving Lives. Protecting People™". A search bar and "Advanced Search" link are visible. Below the header is a green banner with the text "Behavioral Risk Factor Surveillance System". The main content area features a large image of people on a call center line. To the right of the image is the BRFSS logo and a text block: "The Behavioral Risk Factor Surveillance System (BRFSS) is the nation's premier system of health-related telephone surveys that collect state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services. Established in 1984 with 15 states, BRFSS now collects data in all 50 states as well as the District of Columbia and three U.S. territories. BRFSS completes more than 400,000 adult interviews each year, making it the largest continuously conducted health survey system in the world. [See More.](#)" Below the image is a text box: "2018 BRFSS Data Now Available" and "View the latest 2018 BRFSS Annual Data".

A blue banner with a green graphic on the left showing a bar chart. The text on the right reads "National Hospital Discharge Survey".

The **DHS** Program
Demographic and Health Surveys

SEARCH | LOGIN

Select Language



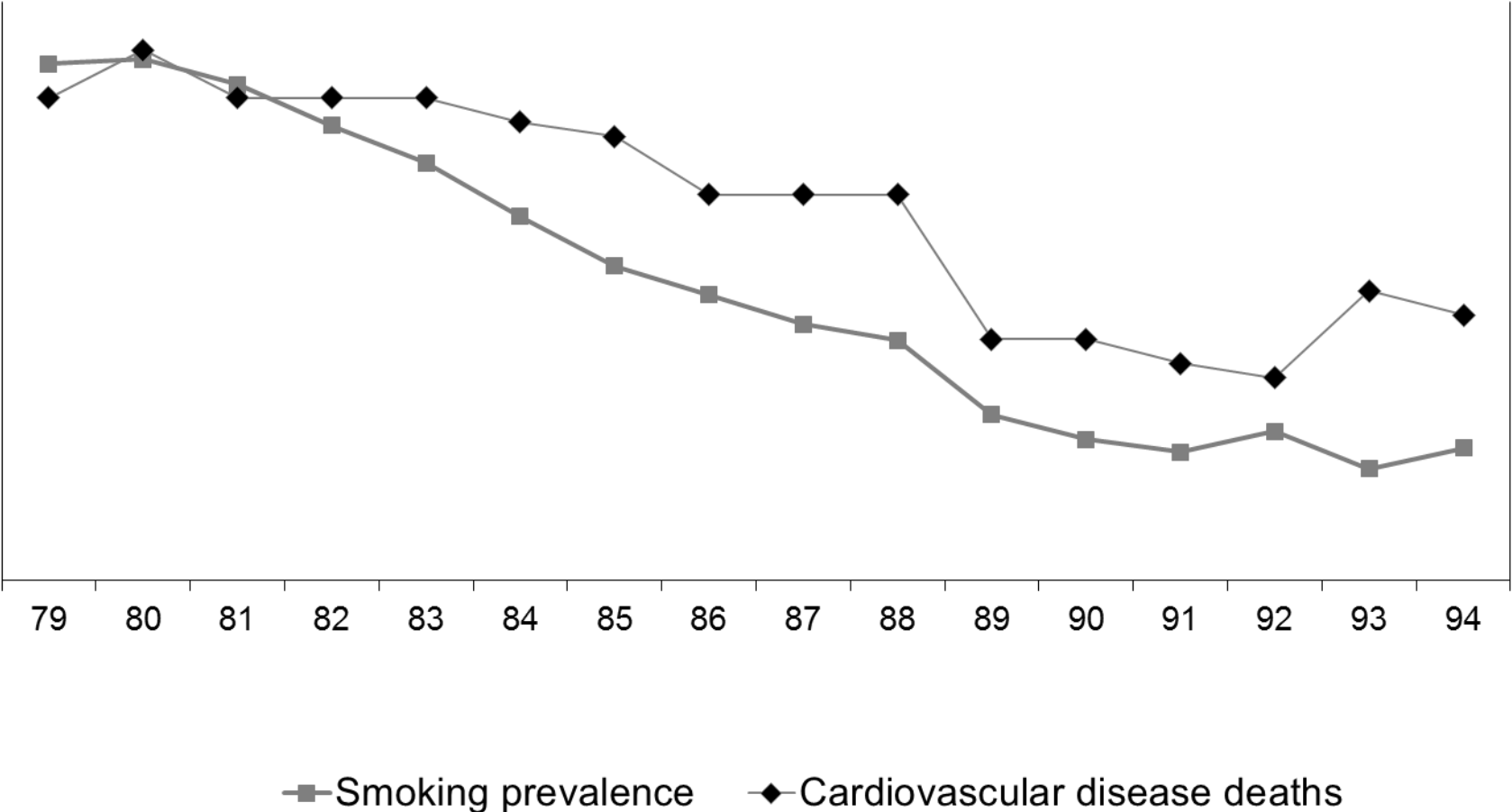
WHO WE ARE | WHAT WE DO | WHERE WE WORK | DATA | PUBLICATIONS | TOPICS

A red slide with a white left arrow and a white right arrow. On the left is a vertical green card with the text "Zambia" at the top, a landscape image of a sunset over a field, and "Demographic and Health Survey 2018" at the bottom. To the right of the card is the text: "In Zambia, 84% of births are delivered in a health facility." Below this is "2018 Zambia DHS »". At the bottom center are five small white dots, with the first one being larger.

Ecologic Studies

- Studies in which the units of analysis are populations or groups of people, rather than individuals
- Useful for hypothesis generation

Cardiovascular Disease Deaths and Smoking Prevalence (Males, 1979-1994)

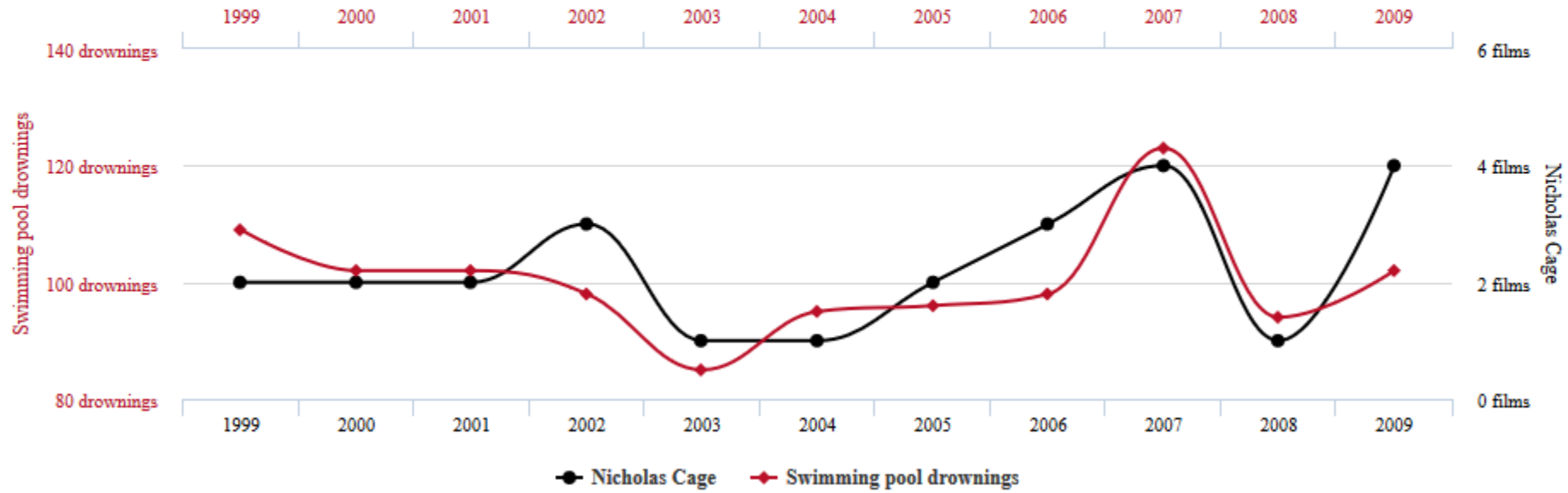


Ecologic Fallacy

- Each individual in the population is characterized by the average for the population
- Bias may occur because an association observed between variables on an aggregate level does not necessarily represent the association that exists at an individual level
 - Because you don't know the joint distribution of exposure/disease/other factors at an individual level

Number of people who drowned by falling into a pool correlates with Films Nicolas Cage appeared in

Correlation: 66.6% (r=0.666004)



Data sources: Centers for Disease Control & Prevention and Internet Movie Database

tylervigen.com

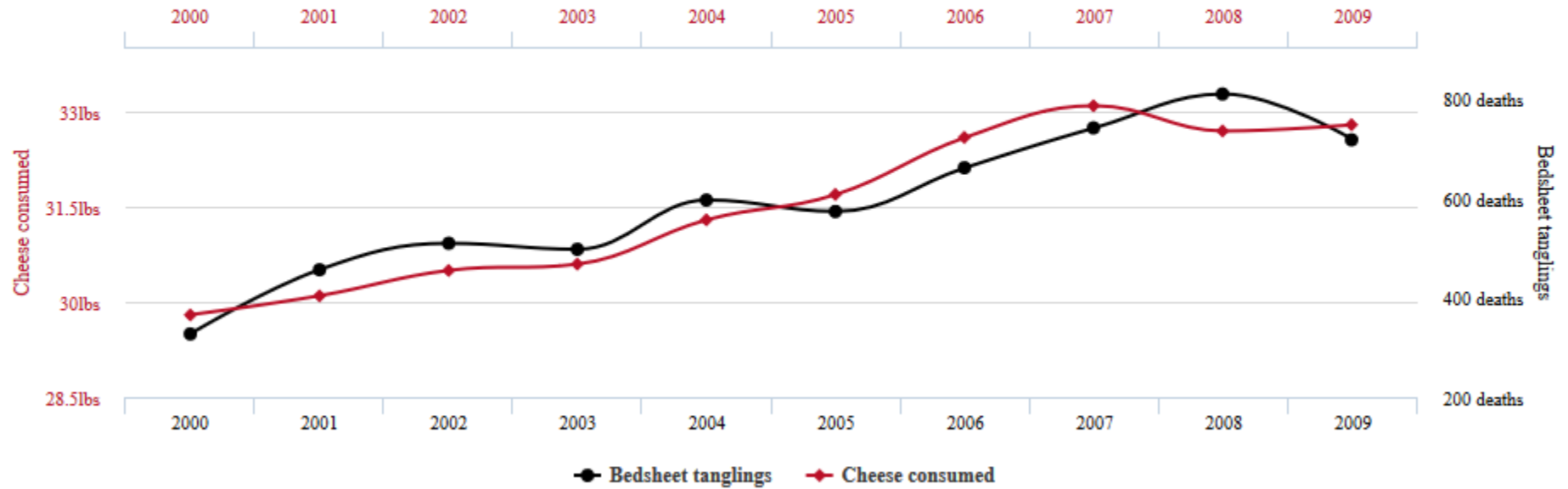
<https://www.tylervigen.com/spurious-correlations>

Per capita cheese consumption

correlates with

Number of people who died by becoming tangled in their bedsheets

Correlation: 94.71% (r=0.947091)



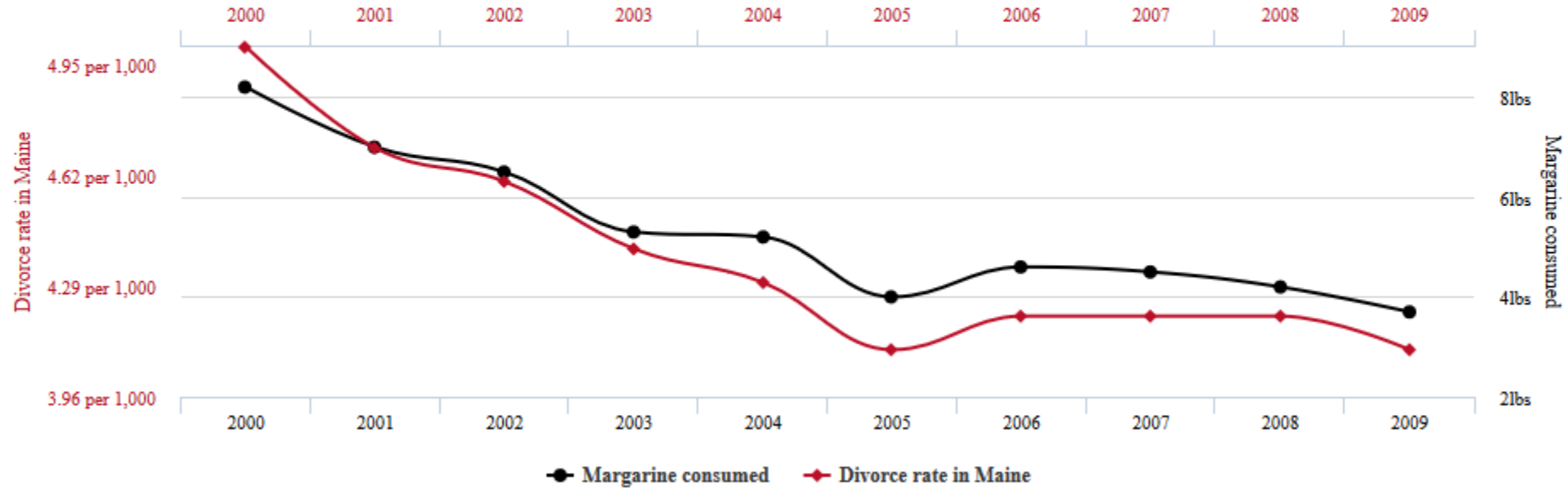
tylervigen.com

Data sources: U.S. Department of Agriculture and Centers for Disease Control & Prevention

<https://www.tylervigen.com/spurious-correlations>

Divorce rate in Maine correlates with Per capita consumption of margarine

Correlation: 99.26% ($r=0.992558$)



Data sources: National Vital Statistics Reports and U.S. Department of Agriculture

tylervigen.com

<https://www.tylervigen.com/spurious-correlations>

To view at your leisure

- https://www.youtube.com/watch?time_continue=6&v=jbkSRLYSojo&feature=emb_logo
- Hans Rosling's 200 Countries, 200 Years, 4 Minutes – The Joy of Stats

Case Series

- Studies without a comparison group
- All study subjects have the disease (or the exposure)
- Impossible to make inferences about causality
- Usually the first report of a new disease/syndrome
 - HIV, microcephaly due to Zika, SARS-CoV-2

Example

- 30% of a series of CHD patients are found to be smokers
- Can we conclude that there is an association between CHD and smoking?

Examples of Case Series

Epidemiologic Notes and Reports

***Pneumocystis* Pneumonia --- Los Angeles**

In the period October 1980-May 1981, 5 young men, all active homosexuals, were treated for biopsy-confirmed *Pneumocystis carinii* pneumonia at 3 different hospitals in Los Angeles, California. Two of the patients died. All 5 patients had laboratory-confirmed previous or current cytomegalovirus (CMV) infection and candidal mucosal infection. Case reports of these patients follow.

Patient 1: A previously healthy 33-year-old man developed *P. carinii* pneumonia and oral mucosal candidiasis in March 1981 after a 2-month history of fever associated with elevated liver enzymes, leukopenia, and CMV viremia. The serum complement-fixation CMV titer in October 1980 was 256; in May 1981 it was 32.* The patient's condition deteriorated despite courses of treatment with trimethoprim-sulfamethoxazole (TMP/SMX), pentamidine, and acyclovir. He died May 3, and postmortem examination showed residual *P. carinii* and CMV pneumonia, but no evidence of neoplasia.

Patient 2: A previously healthy 30-year-old man developed *p. carinii* pneumonia in April 1981 after a 5-month history of fever each day and of elevated liver-function tests, CMV viremia, and documented seroconversion to CMV, i.e., an acute-phase titer of 16 and a convalescent-phase titer of 28* in anticomplement immunofluorescence tests. Other features of his illness included leukopenia and mucosal candidiasis. His pneumonia responded to a course of intravenous TMP/SMX, but, as of the latest reports, he continues to have a fever each day.

Patient 3: A 30-year-old man was well until January 1981 when he developed esophageal and oral candidiasis that responded to Amphotericin B treatment. He was hospitalized in February 1981 for *P. carinii* pneumonia that responded to TMP/SMX. His esophageal candidiasis recurred after the pneumonia was diagnosed, and he was again given Amphotericin B. The CMV complement-fixation titer in March 1981 was 8. Material from an esophageal biopsy was positive for CMV.

Patient 4: A 29-year-old man developed *P. carinii* pneumonia in February 1981. He had had Hodgkins disease 3 years earlier, but had been successfully treated with radiation therapy alone. He did not improve after being given intravenous TMP/SMX and corticosteroids and died in March. Postmortem examination showed no evidence of Hodgkins disease, but *P. carinii* and CMV were found in lung tissue.

Patient 5: A previously healthy 36-year-old man with clinically diagnosed CMV infection in September 1980 was seen in April 1981 because of a 4-month history of fever, dyspnea, and cough. On admission he was found to have *P. carinii* pneumonia, oral candidiasis, and CMV retinitis. A complement-fixation CMV titer in April 1981 was 128. The patient has been treated with 2 short courses of TMP/SMX that have been limited because of a sulfa-induced neutropenia. He is being treated for candidiasis with topical nystatin.

JOURNAL ARTICLE

Preliminary Report of Microcephaly Potentially Associated with Zika Virus Infection During Pregnancy — Colombia, January–November 2016

Esther Liliana Cuevas, Van T. Tong, Nathaly Rozo, Diana Valencia, Oscar Pacheco, Suzanne M. Gilboa, Marcela Mercado, Christina M. Renquist, Maritza González, Elizabeth C. Ailes, Carolina Duarte, Valerie Godoshian, Christina L. Sancken, Angelica Maria Rico Turca, Dinorah L. Calles, Martha Ayala, Paula Morgan, Erika Natalia Tolosa Perez, Hernan Quijada Bonilla, Ruben Caceres Gomez, Ana Carolina Estupiñan, Maria Luz Gunturiz, Dana Meaney-Delman, Denise J. Jamieson, Margaret A. Honein and Martha Lucia Ospina Martínez

Morbidity and Mortality Weekly Report
Vol. 65, No. 49 (December 16, 2016), pp.
1409-1413

Research

Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series

BMJ 2020 ; 368 doi: <https://doi.org/10.1136/bmj.m606> (Published 19 February 2020)

Cite this as: *BMJ* 2020;368:m606

Bias

- Deviation of results or inferences from the “truth”
- *Antonym: Validity*

Selection Bias	Information Bias
Completing risks	Differential misclassification
Healthcare access bias	Non-differential misclassification
Length-bias	Detection bias
Neyman bias (incidence/prevalence)	Observer/interviewer bias
Berkson's bias (probability of hospitalization)	Recall bias
Friend control bias	Reporting bias
Citation bias	Hawthorne effect
Publication bias	Lead time bias
Losses/withdrawals to follow up	Will Rogers phenomenon
Missing information	Protopathic bias
Non-response bias	Work up bias (verification bias)
Healthy worker effect	Temporal ambiguity

Bias

- Selection bias
 - Systematic error introduced when the study population does not represent the target population
 - The relationship observed within your study population differs from the relationship among those who didn't make it into your study
- Information bias
 - Generally occurs during data collection
 - An issue of misclassification where an "exposed" person is classified as "unexposed" or a person with the outcome is classified as not having the outcome, or vice versa

Selection Bias

- Distortion in study results due to the manner in which subjects are selected for the study

Examples of Selection Bias

- Bias related to nonresponse
- Bias related to loss to follow-up

Nonresponse

- Nonresponse may be due to refusal, migration, death, missing records
- Nonrespondents may differ from respondents

Nonresponse

Example:

- Subjects who refuse to participate in a study of smoking and CHD may be more likely to be smokers

Loss to Follow-Up

- In cohort studies and randomized controlled trials, persons who are lost to follow-up may differ from those who remain in the study

Loss to Follow-Up

Example:

- Prospective cohort study of the effect of smoking on CHD
- Study dropouts may be more likely to be smokers

What Can Be Done?

- Be aware of potential sources of selection bias
- Proper study design

Information Bias

- Errors in classification of subjects with respect to disease or exposure

Information Bias

Example:

- Case-control study of CHD and smoking
- Persons with CHD may be more likely to deny smoking history

What Can Be Done?

- Use data collection tools that have been validated, pretested
- Use similar data collection methods for all subjects in study (cases/controls, exposed/unexposed)
- Ensure that research staff are “blind” to subjects’ disease and exposure status

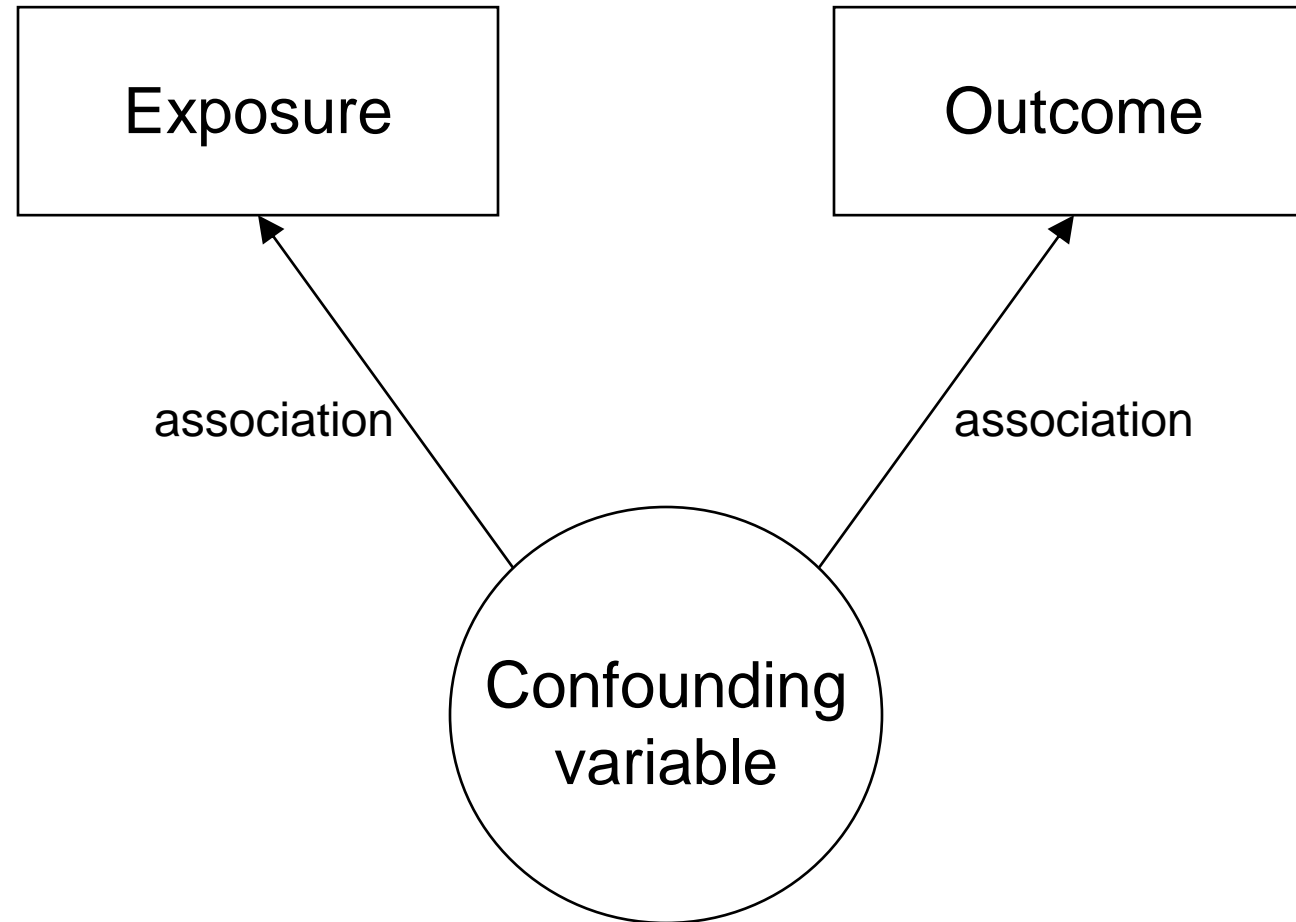
Confounding

- Confounding is the distortion of an exposure-outcome association brought about by the association of another factor with both outcome and exposure
- A confounder is a variable that masks the true relationship between an exposure and a disease

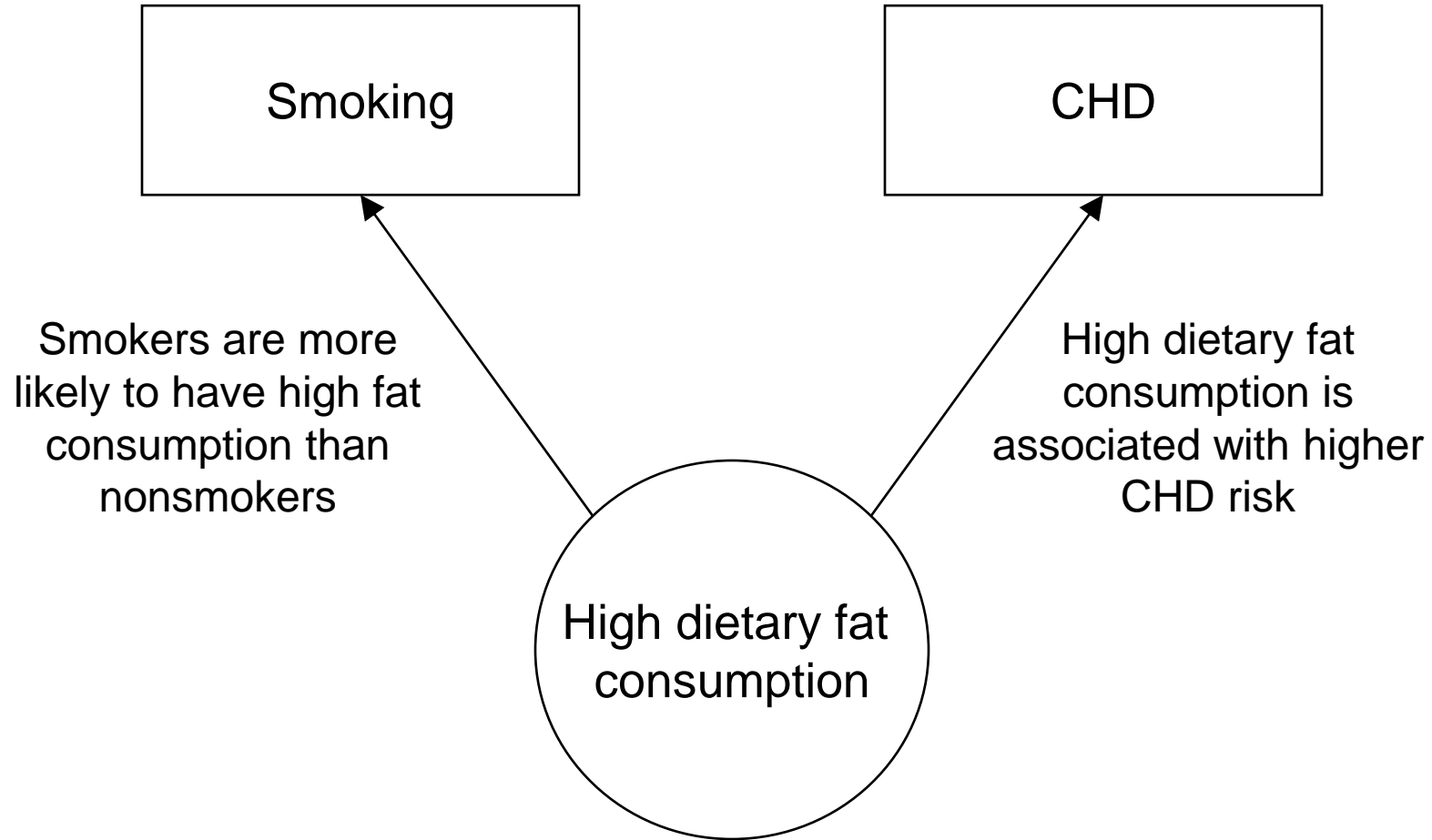
Confounding

- In order for confounding to occur, a variable must be a risk factor for the disease and be distributed differently among exposed and nonexposed
- If only one of these conditions is met, there will be no confounding

Confounding



Example



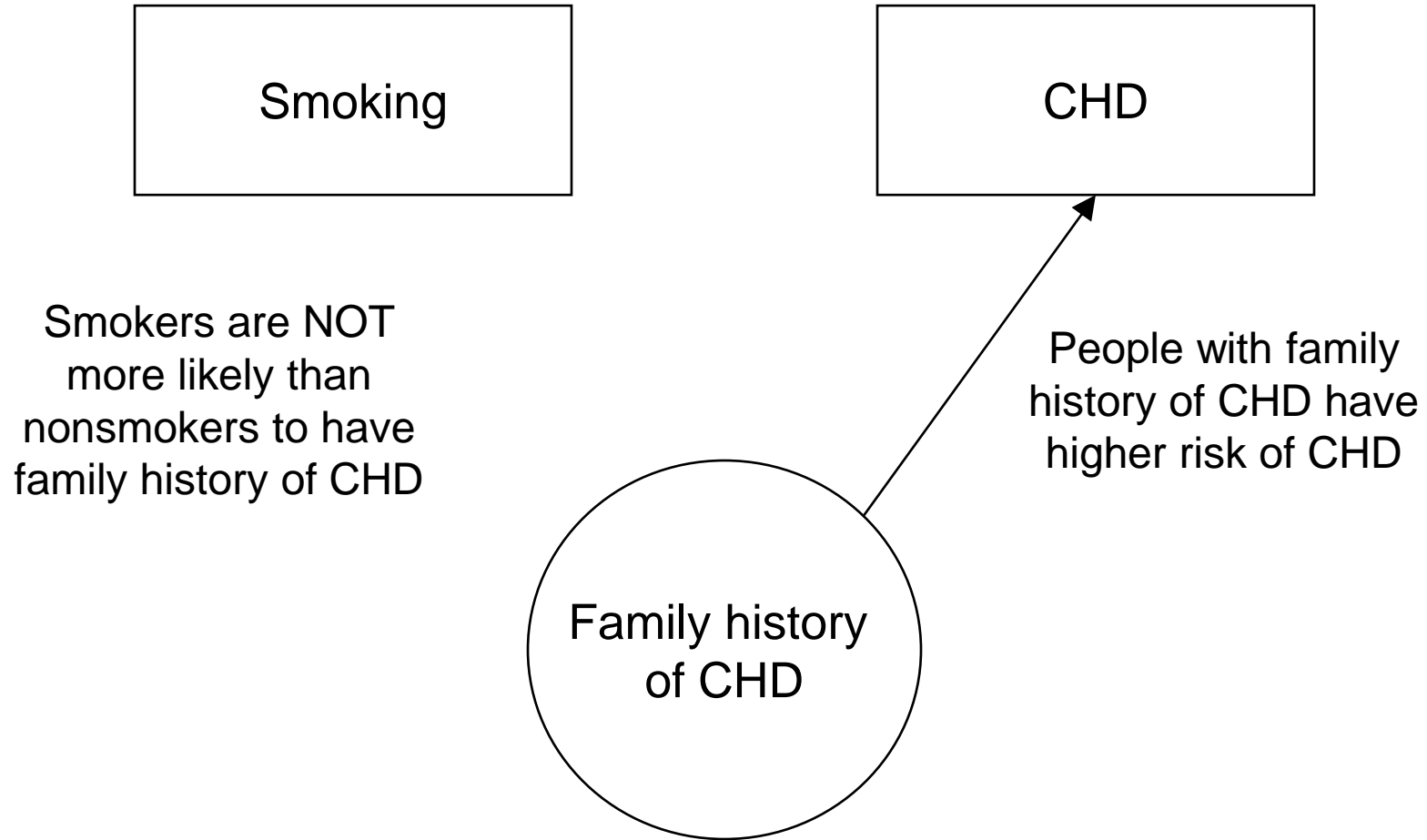
Example

- Suppose you wish to study the effect of smoking on the risk of CHD
- Smokers are more likely to have high dietary fat consumption than nonsmokers
- High dietary fat consumption is a risk factor for CHD
- Therefore, high dietary fat consumption is a confounder

Example

- Suppose you wish to study the effect of smoking on the risk of CHD
- Family history of CHD is a risk factor for CHD
- Family history of CHD is not more common in smokers than nonsmokers
- Therefore, family history of CHD is not a confounder

Example



Control of Confounding

- If a variable is a confounder, then controlling for that variable will result in a change in the estimated effect of the exposure on the disease

Control of Confounding

At design stage:

- Randomization
- Matching
- Restricting study to certain groups

At analysis stage:

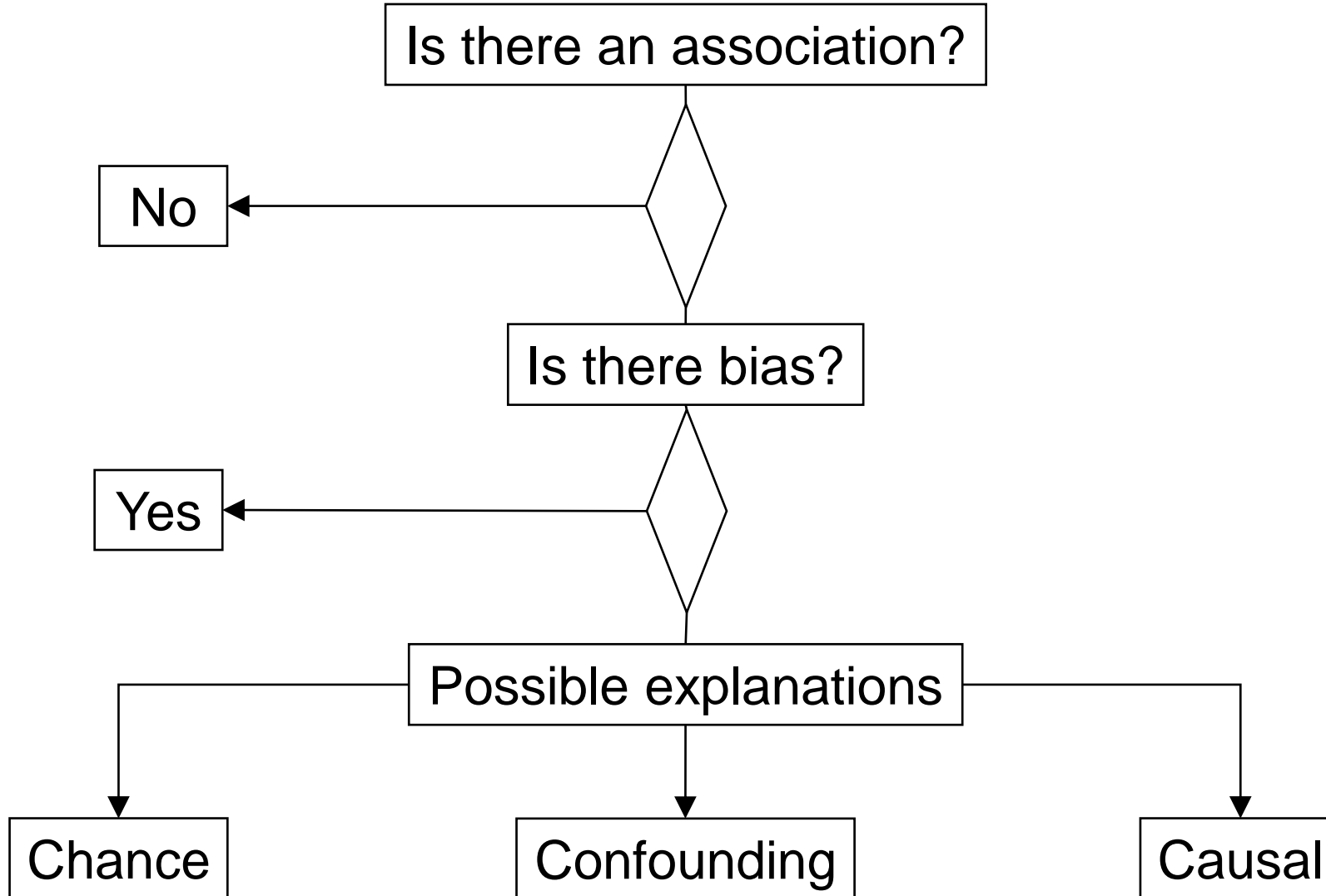
- Statistical methods (stratification, standardization, regression)

Why Is Confounding Important?

- Interferes with search for causal associations
- If association is not causal, intervention will not be effective

	Cross-sectional	Case-control	Cohort	Clinical trial
Selection bias:				
• <i>Nonresponse</i>	×	×	×	×
• <i>Loss to follow-up</i>			×	×
Information bias	×	×	×	×
Confounding	×	×	×	

Epidemiologic Reasoning



Criteria for Causality

Temporality*

- The cause must precede the effect in time

Strength of the association*

- Strong associations are more likely to be causal than weak associations

Dose-response effect*

- If higher levels of exposure result in higher risk of disease, the association is more likely to be causal

Consistency

- Repeated observation of the association in different populations under different circumstances supports causality

Biological plausibility

- Causality is supported if the association makes sense in the context of current biological knowledge

** Applied to findings of a single study*

Summary of Common Study Designs

Design	Advantages	Disadvantages
Case Control	<ul style="list-style-type: none">CheaperQuicker/easier to conductGood for long latencyCan assess multiple exposuresGood for rare diseases	<ul style="list-style-type: none">Prone to bias, including selectionRetrospective, prone to recall biasTypically, only assess one outcomeCannot establish riskCannot establish prevalence
Cohort	<ul style="list-style-type: none">ProspectiveCan directly establish riskCan assess multiple outcomesGood for rare exposures	<ul style="list-style-type: none">Prone to bias, including selectionMore expensiveLonger/harder to conductNot good for rare diseasesNot good with long latency periods
Clinical Trials	<ul style="list-style-type: none">ProspectiveCan directly establish riskEliminates selection bias	<ul style="list-style-type: none">More expensiveHarder to conductPossible ethical issues

Useful tools for study design and evaluation

- CONSORT (RCTs)

www.consort-statement.org/

- STROBE (observational studies)

<https://www.strobe-statement.org/index.php?id=strobe-home>

- Quasi-experimental

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5669452/>

Name that study design!

- 1000 UMMC patients are enrolled and assessed for ETOH use. They are then followed for 10 years to see if they develop esophageal cancer.
- 30 patients with esophageal cancer at the VA are enrolled and compared with 30 VA patients without esophageal cancer to determine what factors are associated with this type of cancer.
- 30 Internal Medicine Residents are randomly assigned to either review the medical literature with the support of an Epidemiologist or review the medical literature alone and then their desire to ever work with an Epidemiologist again is assessed.

I CAN'T BELIEVE SCHOOLS
ARE STILL TEACHING KIDS
ABOUT THE NULL HYPOTHESIS.

I
I REMEMBER READING A BIG
STUDY THAT CONCLUSIVELY
DISPROVED IT *YEARS* AGO.

