Selection Bias in Epidemiological Studies

Madhukar Pai, MD, PhD
Assistant Professor
Department of Epidemiology & Biostatistics
McGill University, Montreal, Canada
Email: madhukar.pai@mcgill.ca

INDIA TODAY

HEALTH

April 27, 1998

INDIA TODAY

- POLITICS
- BUSINESS
- ENTERTAINMENT &THE ARTS
- PEOPLE
- ABOUT US

COVER STORY

Rising Impotency

A silent, embarrassing affliction sweeps Indian bedrooms, with men from all age groups and social classes desperately turning to doctors for cures.

By Madhu Jain, Subhadra Menon and Ramesh Vinayak

Lalit Bhasin is the kind of man most men envy. Just a year ago he would have thought so too. At 39, the suave, successful banker had a well-turned out wife and two precocious children, swung a mean club on the golf course. But a few months ago, things began to go wrong in the bedroom. The worst thing that could happen to a man happened to him: he became impotent. "It drove me nuts," confesses Bhasin. "I had money, but without my virility I was only half a man."

Leading Indian News Magazine:
One out of every 10 Indian males could be impotent!

India Today, 1998

INDIA TODAY

HEALTH

April 27, 1998

INDIA TODAY

- POLITICS
- BUSINESS
- ENTERTAINMENT &THE ARTS
- PEOPLE
- ABOUT US

COVER STORY

Rising Impotency

A silent, embarrassing affliction sweeps Indian bedrooms, with men from all age groups and social classes desperately turning to doctors for cures.

By Madhu Jain, Subhadra Menon and Ramesh Vinayak

Lalit Bhasin is the kind of man most men envy. Just a year ago he would have thought so too. At 39, the suave, successful banker had a well-turned out wife and two precocious children, swung a mean club on the golf course. But a few months ago, things began to go wrong in the bedroom. The worst thing that could happen to a man happened to him: he became impotent. "It drove me nuts," confesses Bhasin. "I had money, but without my virility I was only half a man."

He didn't see a connection but whistle-stop work schedules and jetting round the globe chasing deals and dreams had sapped his sexual vitality. And so began a nightmarish odyssey: rounds of quacks and charlatans. It took Rs 25,000 and a great desperation before Bhasin finally summoned courage to seek medical help.

Bhasin's ordeal is not unique. In fact, it has become particularly common. The proverbial headache is getting to be the male preserve. One out of every 10 Indian males could be impotent, according to a survey of 1,500 men done in Delhi by the Alpha One Andrology Centre at Aashlok Hospital. "Impotence is a silent epidemic that is sweeping across the nation, the average victim being a middle-aged male otherwise healthy and successful," says the centre's director, urologist Vikram Sharma. Big city

On closer look…

India Today, 1998

Antidepressant medications: do they work?

Growing evidence of selective publication and publication bias in trials of antidepressant drugs

Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S., Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.

ABSTRACT

BACKGROUND

Evidence-based medicine is valuable to the extent that the evidence base is complete and unbiased. Selective publication of clinical trials — and the outcomes within those trials — can lead to unrealistic estimates of drug effectiveness and alter the apparent risk–benefit ratio.

METHODS

We obtained reviews from the Food and Drug Administration (FDA) for studies of 12 antidepressant agents involving 12,564 patients. We conducted a systematic literature search to identify matching publications. For trials that were reported in the literature, we compared the published outcomes with the FDA outcomes. We also compared the effect size derived from the published reports with the effect size derived from the entire FDA data set.

RESULTS

Among 74 FDA-registered studies, 31%, accounting for 3449 study participants, were not published. Whether and how the studies were published were associated with the study outcome. A total of 37 studies viewed by the FDA as having positive results were published; 1 study viewed as positive was not published. Studies viewed by the FDA as having negative or questionable results were, with 3 exceptions, either not published (22 studies) or published in a way that, in our opinion, conveyed a positive outcome (11 studies). According to the published literature, it appeared that 94% of the trials conducted were positive. By contrast, the FDA analysis showed that 51% were positive. Separate meta-analyses of the FDA and journal data sets showed that the increase in effect size ranged from 11 to 69% for individual drugs and was 32% overall.

From the Departments of Psychiatry (E.H.T., A.M.M.) and Pharmacology (E.H.T.), Oregon Health and Science University; and the Behavioral Health and Neurosciences Division, Portland Veterans Affairs Medical Center (E.H.T., A.M.M., R.A.T.) — both in Portland, OR; the Department of Psychology, Kent State University, Kent, OH (E.L.); the Department of Psychology, University of California–Riverside, Riverside (R.R.); and Harvard University, Cambridge, MA (R.R.). Address reprint requests to Dr. Turner at Portland VA Medical Center, P3MHDC, 3710 SW US Veterans Hospital Rd., Portland, OR 97239, or at turnere@ohsu.edu.

N Engl J Med 2008;358:252-60.

OPEN ACCESS Freely available online

PLoS MEDICINE

Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food and Drug Administration

Irving Kirsch[1*], Brett J. Deacon[2], Tania B. Huedo-Medina[3], Alan Scoboria[4], Thomas J. Moore[5], Blair T. Johnson[3]

1 Department of Psychology, University of Hull, Hull, United Kingdom, 2 University of Wyoming, Laramie, Wyoming, United States of America, 3 Center for Health, Intervention, and Prevention, University of Connecticut, Storrs, Connecticut, United States of America, 4 Department of Psychology, University of Windsor, Windsor, Ontario, Canada, 5 Institute for Safe Medication Practices, Huntingdon Valley, Pennsylvania, United States of America

Funding: The authors received no specific funding for this study.

Competing Interests: IK has received consulting fees from Squibb and Pfizer. BJD, TBH, AS, TJM, and BTJ have no competing interests.

Academic Editor: Phillipa Hay, University of Western Sydney, Australia

Citation: Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, et al. (2008) Initial severity and antidepressant benefits: A meta-analysis of data submitted to the Food and Drug Administration. PLoS Med 5(2): e45. doi:10.1371/journal.pmed.0050045

Received: January 23, 2007
Accepted: January 4, 2008
Published: February 26, 2008

Abbreviations: *d*, standardized mean difference; FDA, US Food and Drug Administration; HRSD, Hamilton Rating Scale of Depression; LOCF, last observation carried forward; NICE, National Institute for Clinical Excellence; *SD*, standard deviation of the change score

ABSTRACT

Background

Meta-analyses of antidepressant medications have reported only modest benefits over placebo treatment, and when unpublished trial data are included, the benefit falls below accepted criteria for clinical significance. Yet, the efficacy of the antidepressants may also depend on the severity of initial depression scores. The purpose of this analysis is to establish the relation of baseline severity and antidepressant efficacy using a relevant dataset of published and unpublished clinical trials.

Methods and Findings

We obtained data on all clinical trials submitted to the US Food and Drug Administration (FDA) for the licensing of the four new-generation antidepressants for which full datasets were available. We then used meta-analytic techniques to assess linear and quadratic effects of initial severity on improvement scores for drug and placebo groups and on drug–placebo difference scores. Drug–placebo differences increased as a function of initial severity, rising from virtually no difference at moderate levels of initial depression to a relatively small difference for patients with very severe depression, reaching conventional criteria for clinical significance only for patients at the upper end of the very severely depressed category. Meta-regression analyses indicated that the relation of baseline severity and improvement was curvilinear in drug groups and showed a strong, negative linear component in placebo groups.

Conclusions

Drug–placebo differences in antidepressant efficacy increase as a function of baseline severity, but are relatively small even for severely depressed patients. The relationship between initial severity and antidepressant efficacy is attributable to decreased responsiveness to placebo among very severely depressed patients, rather than to increased responsiveness to medication.

Antidepressant medications: do they work?

The Gazette

21°C
Overcast
Detailed Forecast

earch for in The Gazette find Saturday, September 20, 2008

Pills. Placebos. Depression

AARON DERFEL, The Gazette
Published: 17 hours ago

Leah, a bright 25-year-old graduate student with lots of friends, had just moved from Montreal to a university in a new city, and should have been excited about starting her studies.

Instead, she plunged into a depression that dragged her down for months. A feeling of "incredible sadness" washed over her, all the more puzzling because there seemed to be no justification for it.

- - -

View Larger Image

Mimi Israel, chief of psychiatry at the Douglas Mental Health University Institute, says she has helped countless people with the new generation of antidepressants - SSRIs.

JOHN MAHONEY, THE GAZETTE

Email to a friend
Printer friendly

Font: A A A A

She couldn't sleep. She couldn't read. She couldn't engage in conversation with her boyfriend or family. By last winter, she could no longer climb out of bed in the morning to attend her classes.

"It's like I was almost in chains," Leah recalled. "I had no more control over my depression than if I had the stomach flu."

That's when Leah's family put her in touch with Mimi Israel, chief of psychiatry at the Douglas Mental Health University Institute. Leah (who didn't want her last name published) had been taking 20 milligrams a day of Celexa since the winter, but the antidepressant had failed to change her mood.

Israel upped the dosage to 30 milligrams and combined the drug with psychotherapy. Within three weeks, Leah said, her depression had lifted and "I was feeling almost back to normal. It was almost miraculous.

"Honestly, going from 20 to 30 milligrams sounds crazy. It's a little half pill (of a difference), but I feel that it's saved my life."

Leah's experience - along with those of

"I have countless examples of helping people with these medications. So I'm not ready to throw them out.“ - Psychiatrist

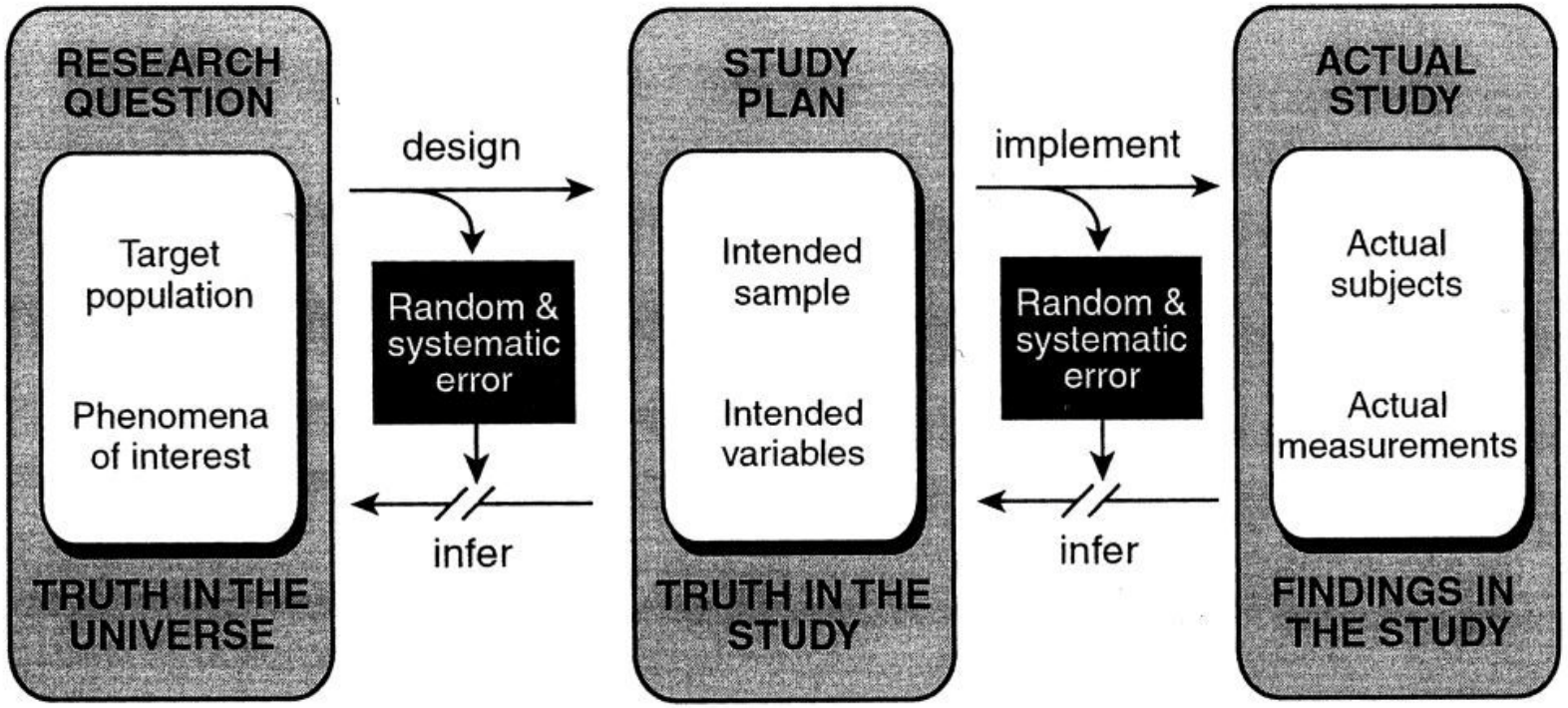
"Well, my personal experience has been that these pills have been extremely effective in the right patients," – Family Physician

So who is right - the authors of those two critical studies or the psychiatrists and GPs on the front lines?

Now lets define selection bias

- "Distortions that result from procedures used to select subjects and from factors that influence participation in the study."
  - Porta M. A dictionary of epidemiology. Oxford, 2008.
- "Error introduced when the study population does not represent the target population"
  - Delgado-Rodriguez et al. J Epidemiol Comm Health 2004
- Defining feature:
  - Selection bias occurs at:
    - the stage of recruitment of participants
    - and/or during the process of retaining them in the study
  - Difficult to correct in the analysis, although one can do sensitivity analyses

Hierarchy of populations



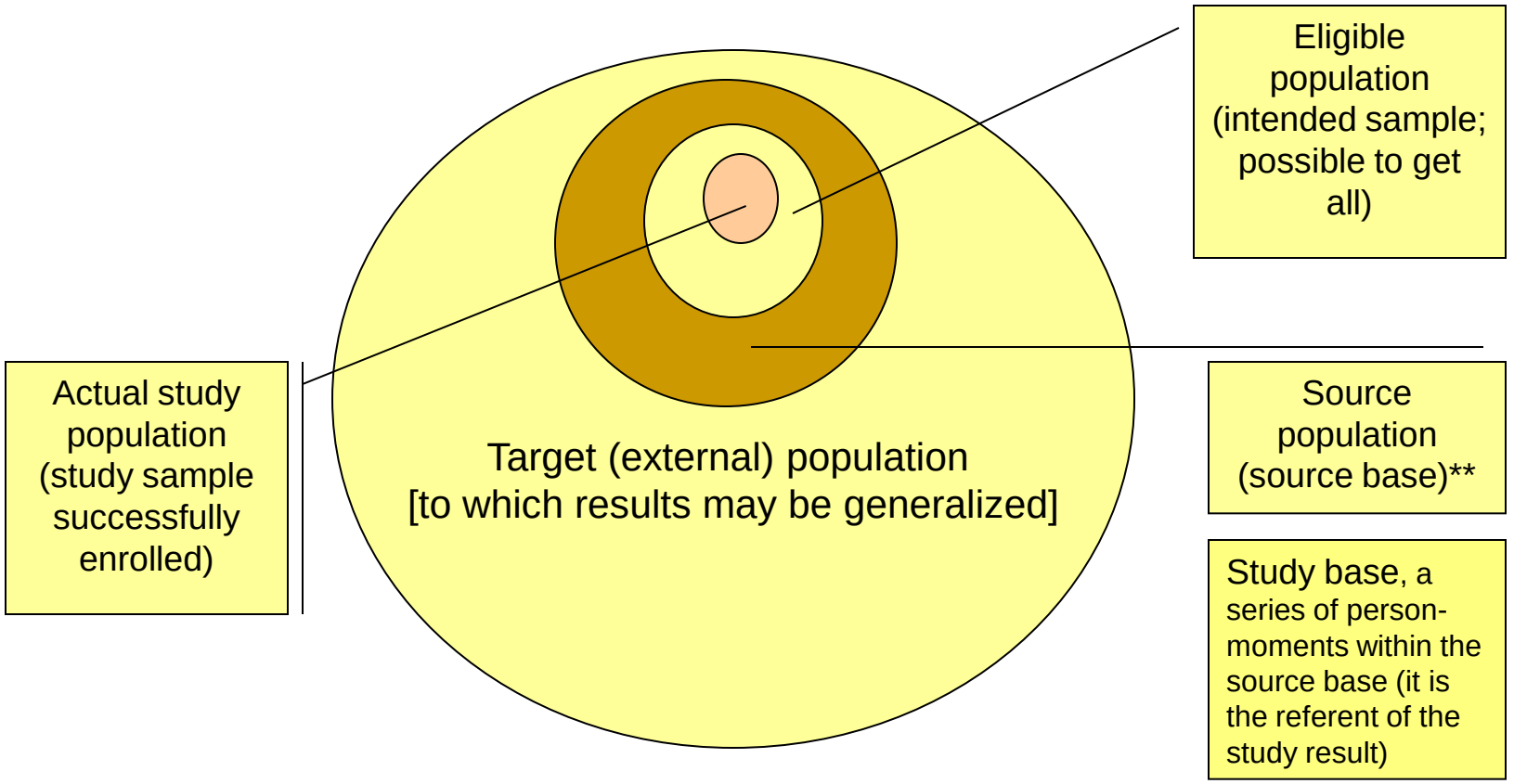
■ **FIGURE 1.6**

Summary of how research works.

Hulley et al. Designing Clinical Research. 2nd Edition. Lippincott Williams & Wilkins, 2001

Hierarchy of populations

Warning: terminology is highly inconsistent! Focus on the concepts, not words!!

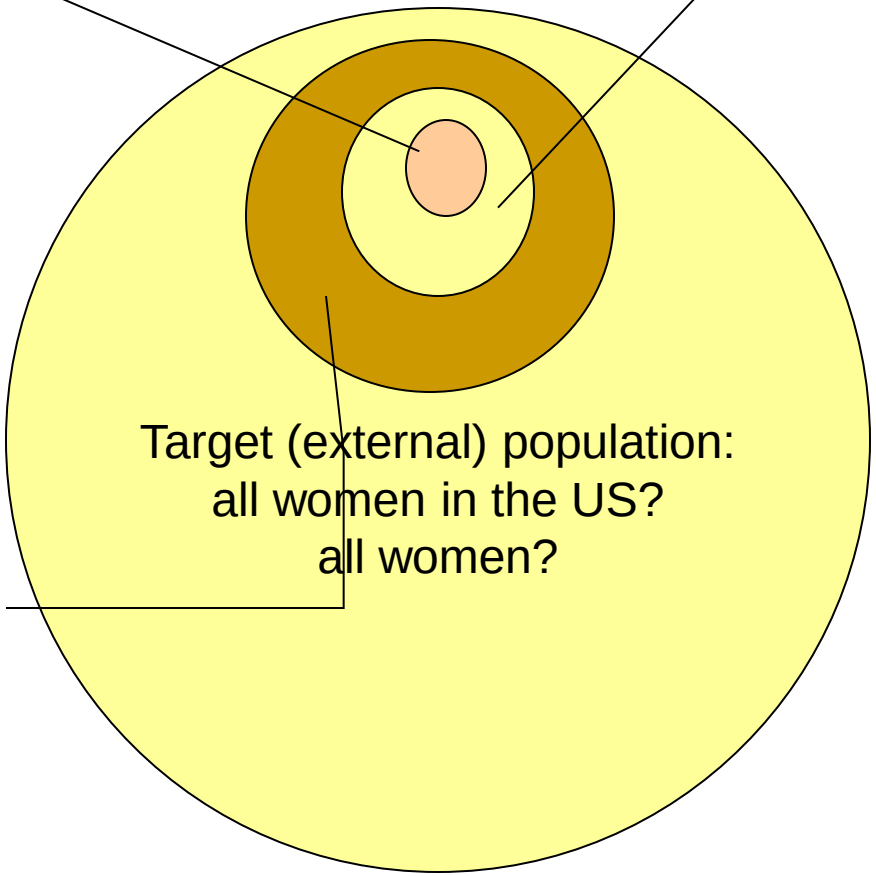


**The source population may be defined directly, as a matter of defining its membership criteria; or the definition may be indirect, as the *catchment population* of a defined way of identifying cases of the illness. The catchment population is, at any given time, the totality of those in the 'were-would' state of: were the illness now to occur, it would be 'caught' by that case identification scheme [Source: Miettinen OS, 2007]

Sampling of populations

Example: Nurses Health Study

Study participants: eligible nurses were enrolled in the cohort if they responded to the baseline questionnaire (122,000 out of 170,000 nurses responded)

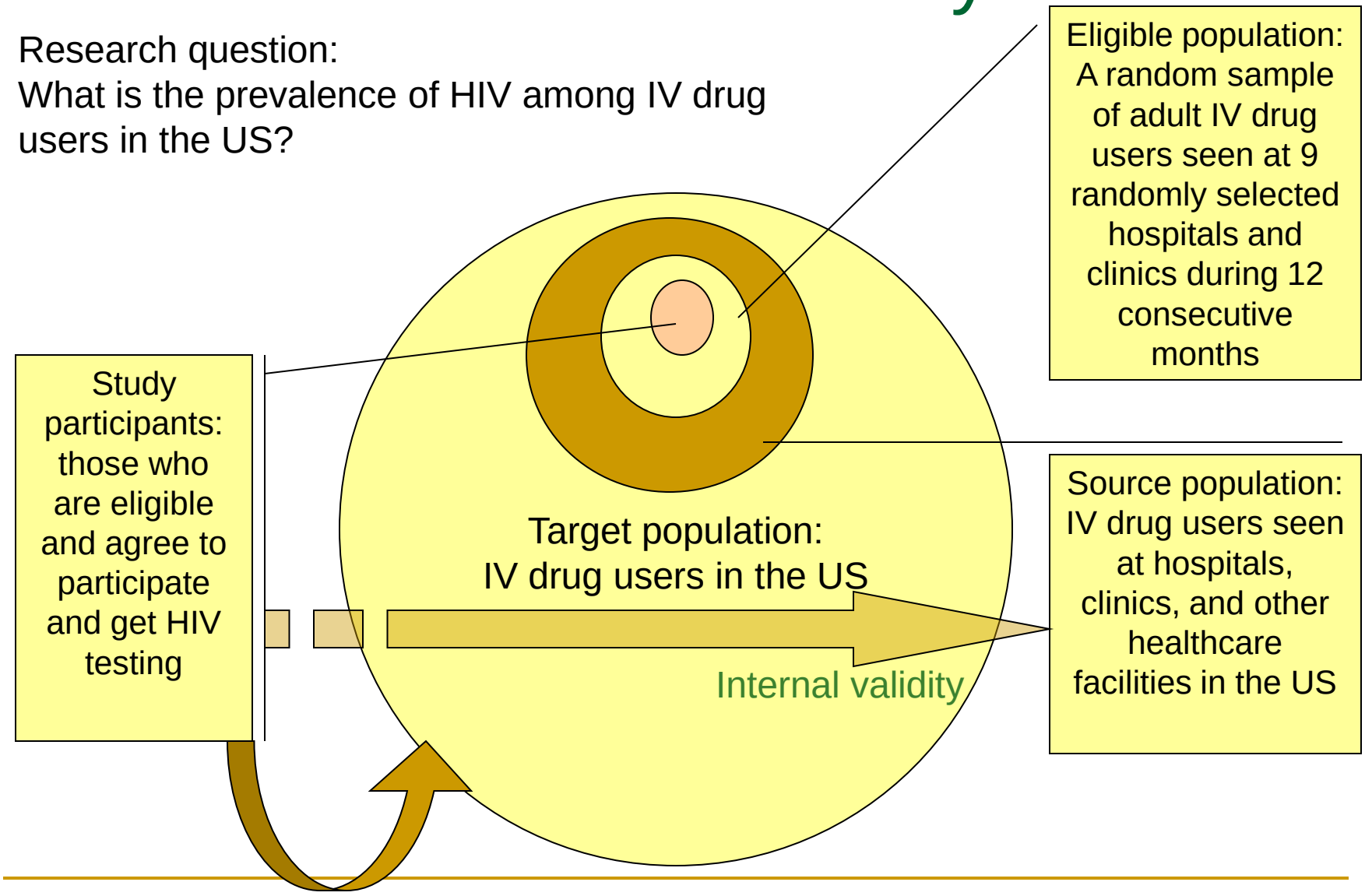


Eligible population: Married, registered nurses who were aged 30 to 55 in 1976, who lived in the 11 most populous states and whose nursing boards agreed to supply the study with their members' names and addresses.

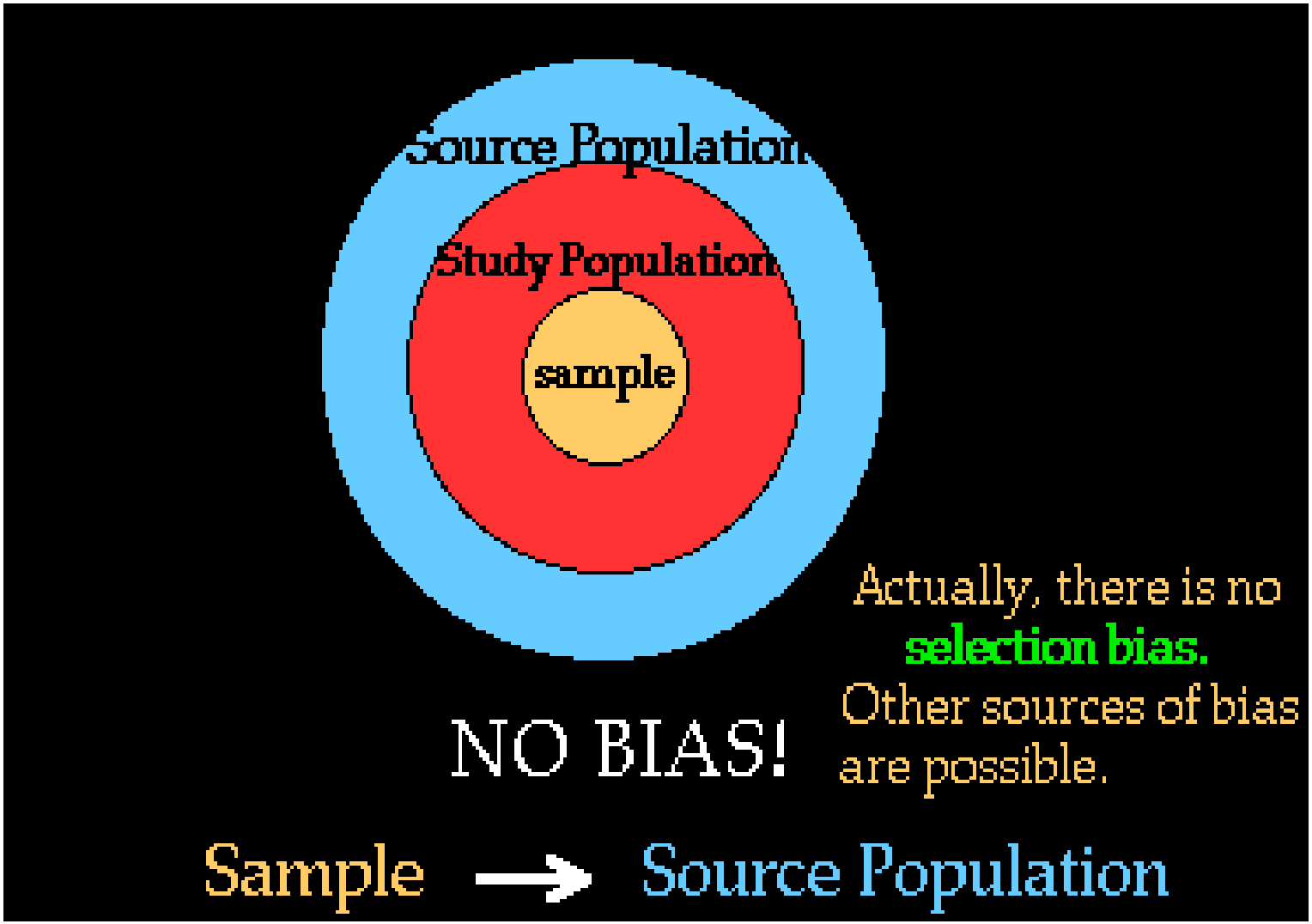
Source population: registered nurses in the US?

Internal vs. External Validity

Research question:
What is the prevalence of HIV among IV drug users in the US?

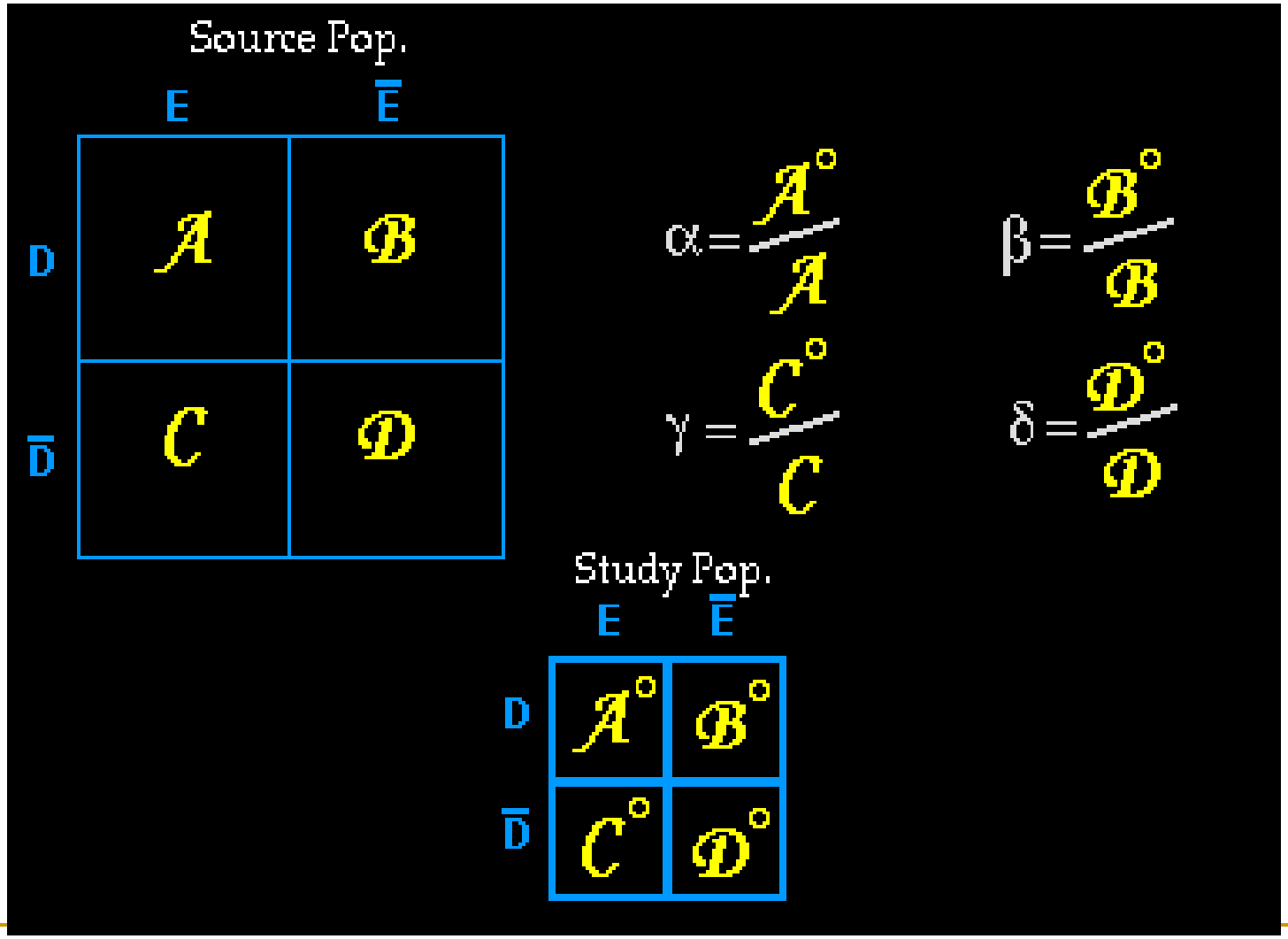


Selection bias can impact both internal and external validity

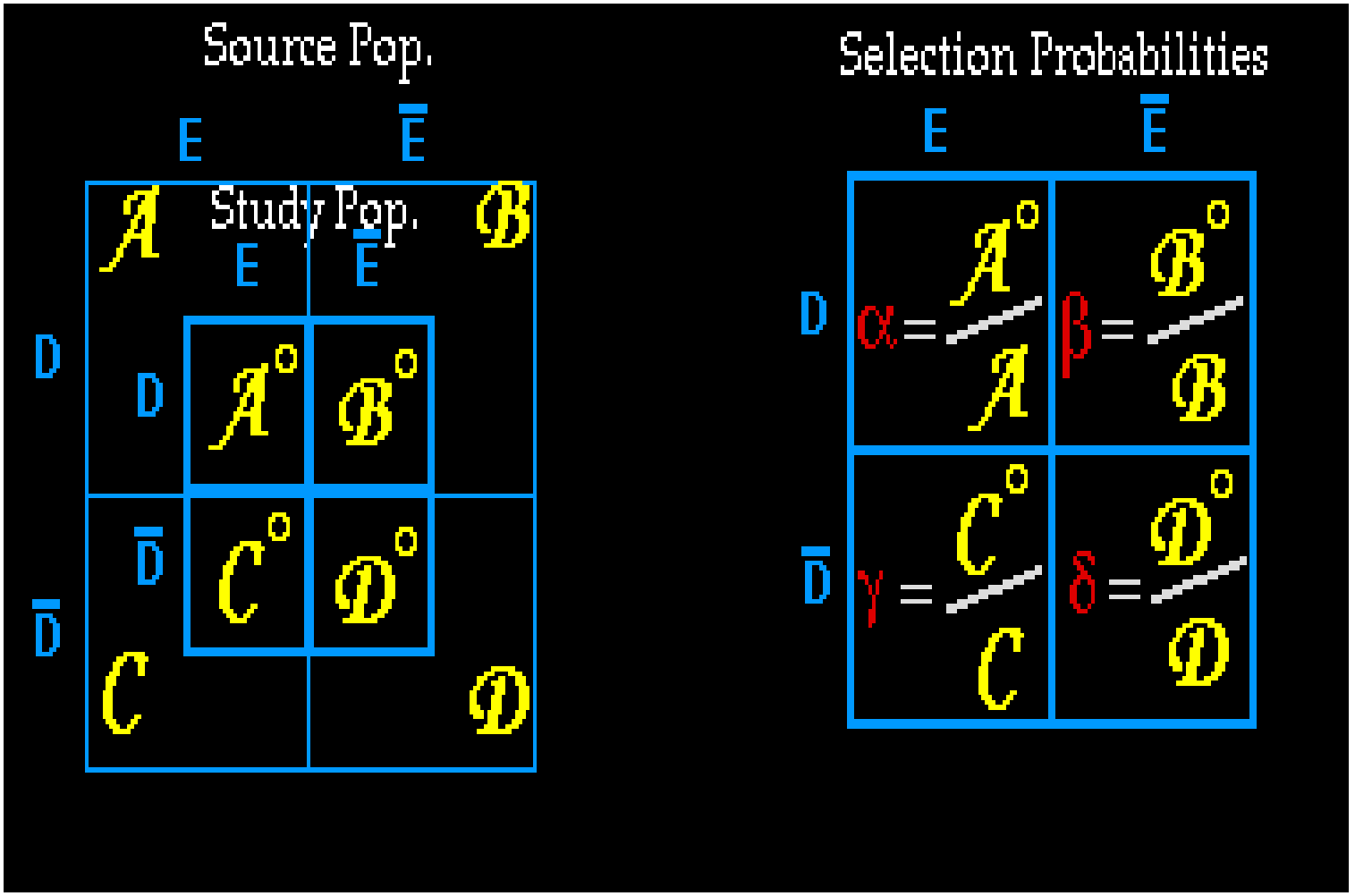


Warning: terminology is highly inconsistent! Focus on the concepts, not words!!

Selection probabilities (also known as ‘sampling fractions’)

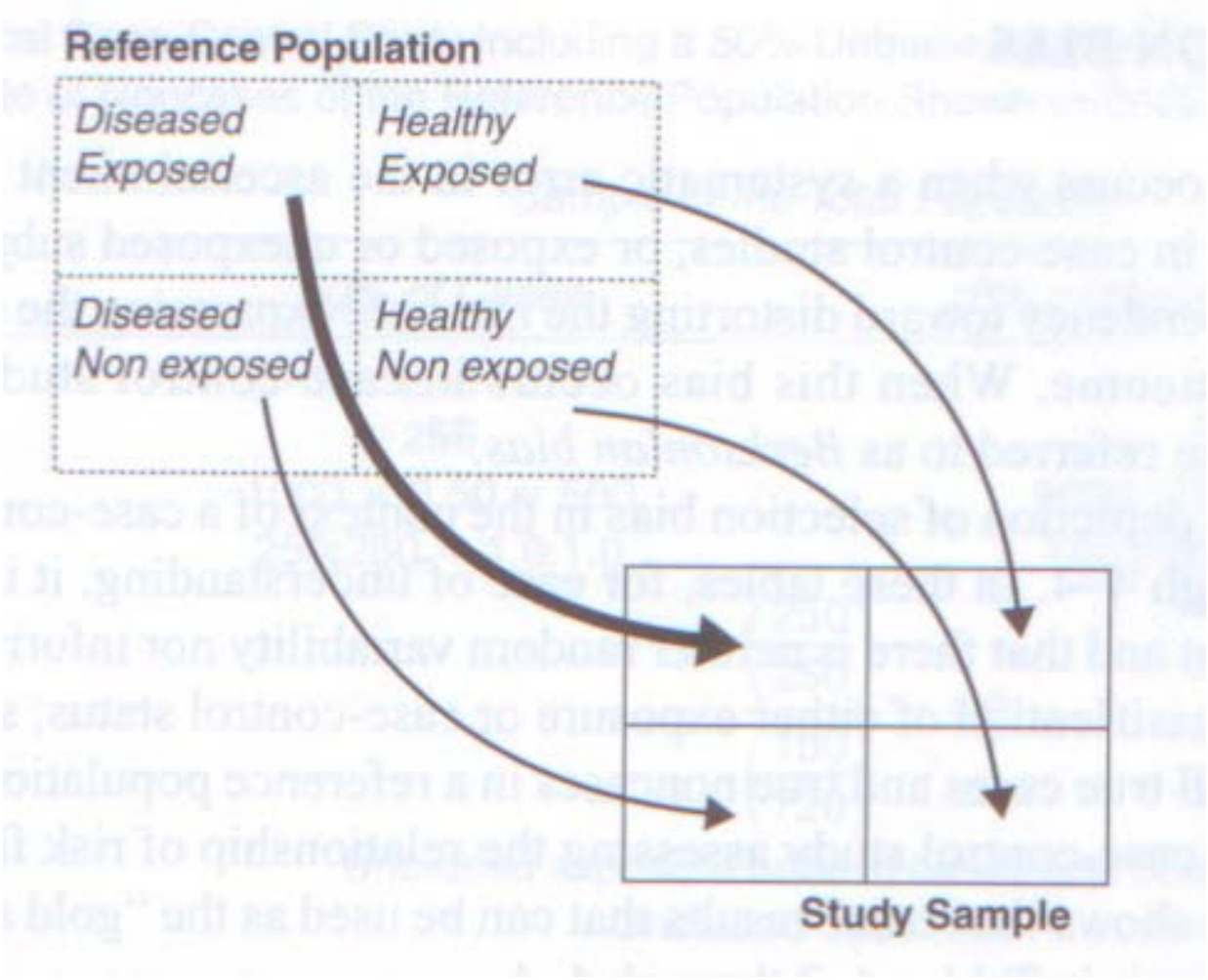


Selection probabilities



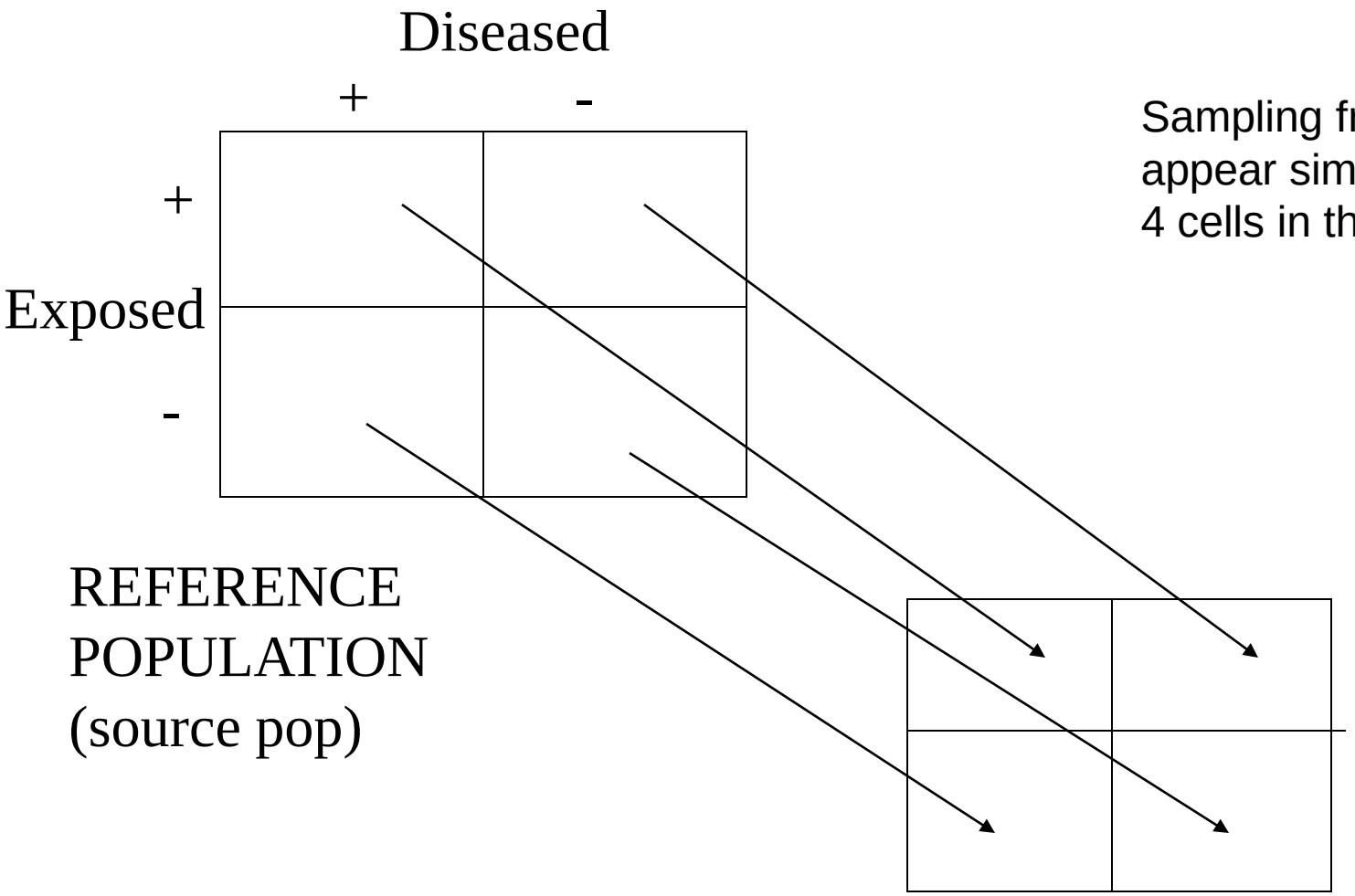
Note: No selection bias if the cross product of α, β, γ, δ = 1

Selection bias occurs when selection probabilities are influenced by exposure or disease status



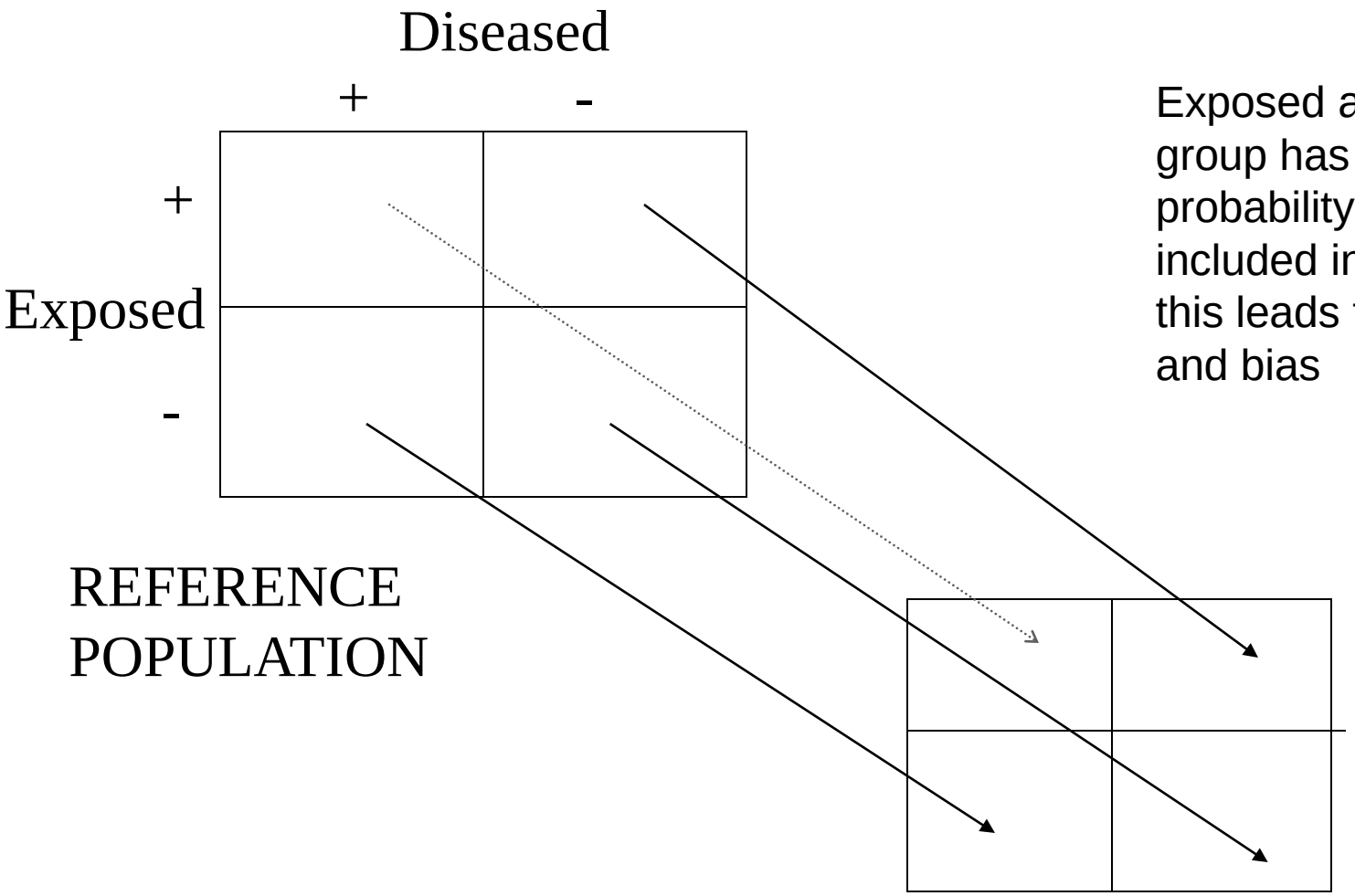
Szklo & Nieto. Epidemiology: Beyond the Basics. 2007

Unbiased Sampling



Sampling fractions appear similar for all 4 cells in the 2 x 2 table

Biased sampling



Exposed and diseased group has a lower probability of being included in the study: this leads to imbalance and bias

Selection bias in randomized controlled trials

- Sources:
  - During randomization (at time t_0)
    - Subversion of randomization due to inadequate concealment of allocation
  - After randomization (during follow up; after time t_0)
    - Attrition***
      - Withdrawals
      - Loss to follow-up
      - Competing risks
      - Protocol violations and “contamination”

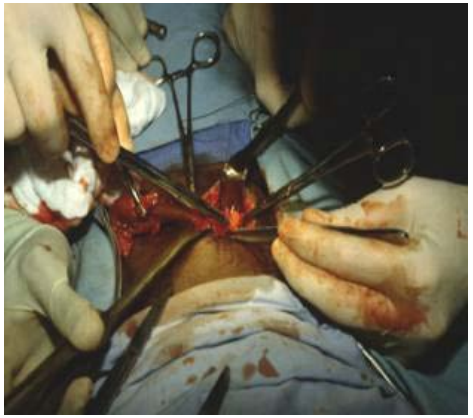
***Also seen in all cohort designs

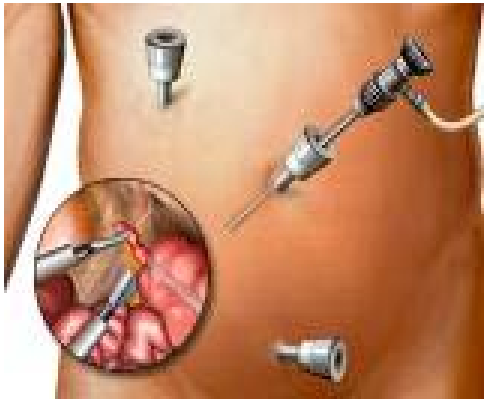
Selection bias in randomized controlled trials

- Examples:
  - Bias due to lack of allocation concealment
    - RCT on thrombolysis with alternating day allocation
    - RCT comparing open versus laparoscopic appendectomy
  - Bias due to attrition
    - RCT comparing medical versus surgical management of cerebrovascular disease

Hansen JB, Smithers BM, Schache D, Wall DR, Miller BJ, Menzies BL. Laparoscopic versus open appendectomy: prospective randomized trial. *World J Surg* 1996;20:17-20; discussion 21.

A prospective randomized trial comparing laparoscopic appendectomy with open appendectomy in patients with a diagnosis of acute appendicitis was conducted between October 1992 and April 1994. Of the 158 patients randomized, 7 patients were excluded because of protocol violations (conversion to laparotomy in 4, appendix not removed in 3). The 151 patients randomized to either a laparoscopic (n = 79) or an open appendectomy (n = 72) showed no difference in sex, age, American Society of Anesthesiology (ASA) rating, or previous abdominal surgery. The histologic classification of normal, catarrhal, inflamed, suppurative, and gangrenous appendicitis was not different between the two groups. Conversion from laparoscopic to open appendectomy was necessary in seven patients (9%) who had advanced forms of appendiceal inflammation. When compared to open appendectomy the laparoscopic group had a longer median operating time (63 minutes versus 40 minutes), fewer wound infections (2% versus 11%), less requirement for narcotic analgesia, and an earlier return to normal activity (median 7 days versus 14 days). There was no difference in morbidity, and both groups had a median time to discharge of 3 days. Laparoscopic appendectomy is as safe as open appendectomy; and despite the longer operating time, the advantages such as fewer wound infections and earlier return to normal activity make it a worthwhile alternative for patients with a clinical diagnosis of acute appendicitis.





Guyatt et al. Users guides to the medical literature. AMA Press, 2002: page 269.

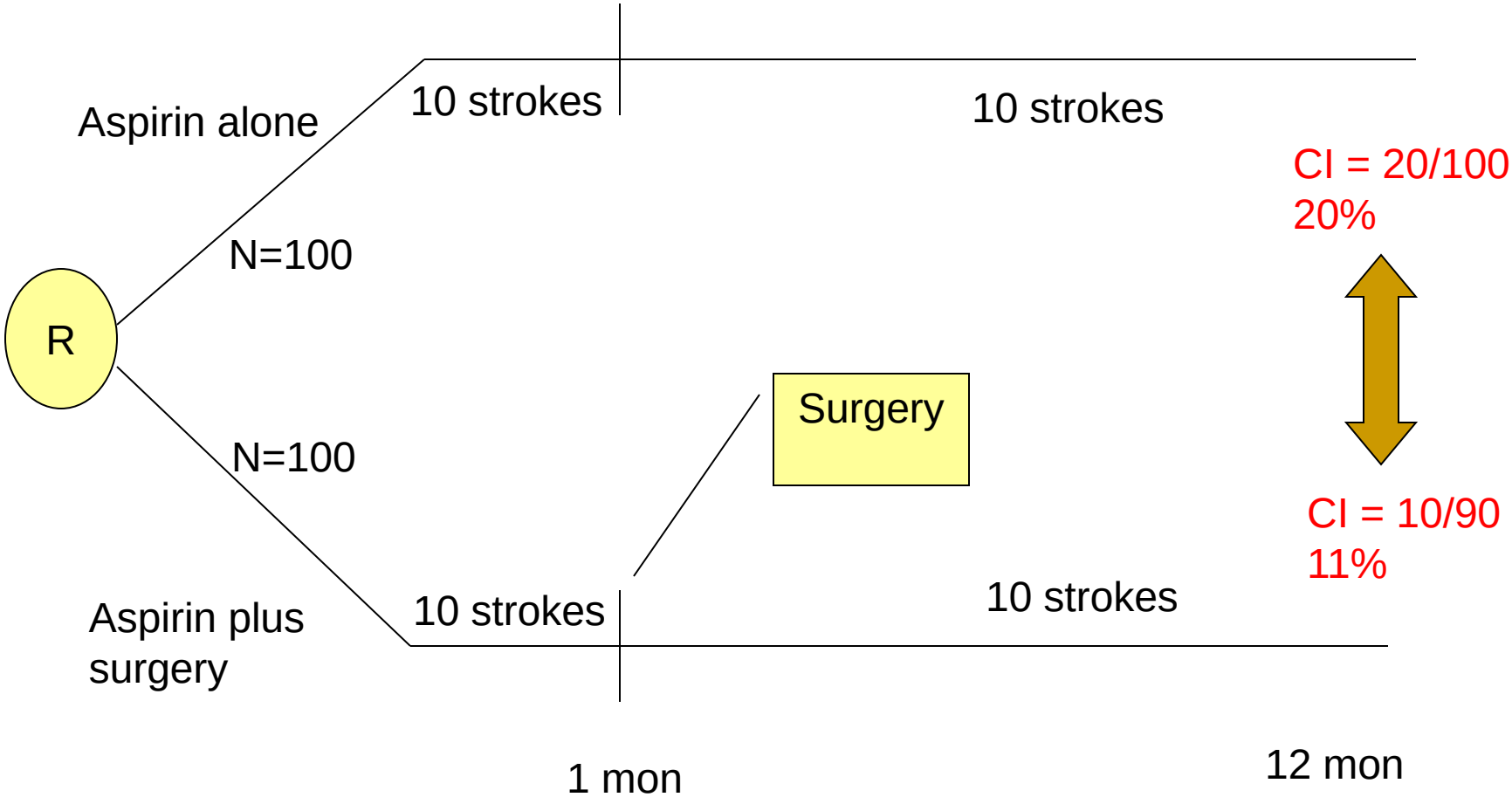
Hansen JB, Smithers BM, Schache D, Wall DR, Miller BJ, Menzies BL. Laparoscopic versus open appendectomy: prospective randomized trial. *World J Surg* 1996;20:17-20; discussion 21.

A prospective randomized trial comparing laparoscopic appendectomy with open appendectomy in patients with a diagnosis of acute appendicitis was conducted between October 1992 and April 1994. Of the 158 patients randomized, 7 patients were excluded because of protocol violations (conversion to laparotomy in 4, appendix not removed in 3). The 151 patients randomized to either a laparoscopic (n = 79) or an open appendectomy (n = 72) showed no difference in sex, age, American Society of Anesthesiology (ASA) rating, or previous abdominal surgery. The histologic classification of normal, catarrhal, inflamed, suppurative, and gangrenous appendicitis was not different between the two groups. Conversion from laparoscopic to open appendectomy was necessary in seven patients (9%) who had advanced forms of appendiceal inflammation. When compared to open appendectomy the laparoscopic group had a longer median operating time (63 minutes versus 40 minutes), fewer wound infections (2% versus 11%), less requirement for narcotic analgesia, and an earlier return to normal activity (median 7 days versus 14 days). There was no difference in morbidity, and both groups had a median time to discharge of 3 days. Laparoscopic appendectomy is as safe as open appendectomy; and despite the longer operating time, the advantages such as fewer wound infections and earlier return to normal activity make it a worthwhile alternative for patients with a clinical diagnosis of acute appendicitis.

- The trial ran smoothly during the day. At night, however, the attending surgeon's presence was required for the laparoscopic procedure but not the open one; and the limited operating room availability made the longer laparoscopic procedure an annoyance.
- Reluctant to call in a consultant, and particularly reluctant with specific senior colleagues, the residents sometimes adopted a practical solution. When an eligible patient appeared, the residents checked the attending staff and the lineup for the operating room and, depending on the personality of the attending surgeon and the length of the lineup, held the translucent envelopes containing orders up to the light. As soon as they found one that dictated an open procedure, they opened that envelope. The first eligible patient in the morning would then be allocated to a laparoscopic appendectomy group according to the passed-over envelope.
- If patients who presented at night were sicker than those who presented during the day, the residents' behavior would bias the results against the open procedure.
- This story demonstrates that if those making the decision about patient eligibility are aware of the arm of the study to which the patient will be allocated --if randomization is unconcealed (unblinded or unmasked)-- they may systematically enroll sicker-- or less sick-- patients to either treatment or control groups.
- This behavior will defeat the purpose of randomization and the study will yield a biased result.
- Careful investigators will ensure that randomization is concealed, for example, through (a) preparation of blinded medication in a pharmacy, (b) remote randomization, in which the individual recruiting the patient makes a call to a methods center to discover the arm of the study to which the patient is allocated, or (c) ensuring that the envelope containing the code is sealed (sealed, opaque envelope).

Guyatt et al. Users guides to the medical literature. AMA Press, 2002: page 269.

Selection bias after randomization (handled by intention-to-treat analysis)



Guyatt et al. Users guides to the medical literature. AMA Press, 2002: page 269.

Selection bias in cohort studies

- Sources:
  - Bias due to a non-representative “unexposed” group
    - Key question: aside from the exposure status, are the exposed and unexposed groups comparable?
      - Has the unexposed population done its job, i.e. generated disease rates that approximate those that would have been found in the exposed population had they lacked exposure (i.e. counterfactual)?
  - Bias due to non-response
    - More likely if non-response is linked to exposure status (e.g. smokers less likely to respond in a study on smoking and cancer)
  - Bias due to attrition (withdrawals and loss to follow up)
    - Bias will occur if loss to follow-up results in risk for disease in the exposed and/or unexposed groups that are different in the final sample than in the original cohort that was enrolled
    - Bias will occur if those who adhere have a different disease risk than those who drop out or do not adhere (‘compliance bias’)

Healthy User and Healthy Continuer Bias: HRT and CHD

- HRT was shown to reduce coronary heart disease (CHD) in women in several observational studies
- Subsequently, RCTs showed that HRT might actually increase the risk of heart disease in women
- What can possibly explain the discrepancy between observational and interventional studies?
  - Women on HRT in observational studies were more health conscious, thinner, and more physically active, and they had a higher socioeconomic status and better access to health care than women who are not on HRT
  - Self-selection of women into the HRT user group could have generated uncontrollable confounding and lead to "healthy-user bias" in observational studies.
  - Also, individuals who adhere to medication have been found to be healthier than those who do not, which could produce a "compliance bias" [healthy user bias]

Michels et al. Circulation. 2003;107:1830

For a more in-depth analysis of this case study, see B-File #1

Selection bias in cohort studies

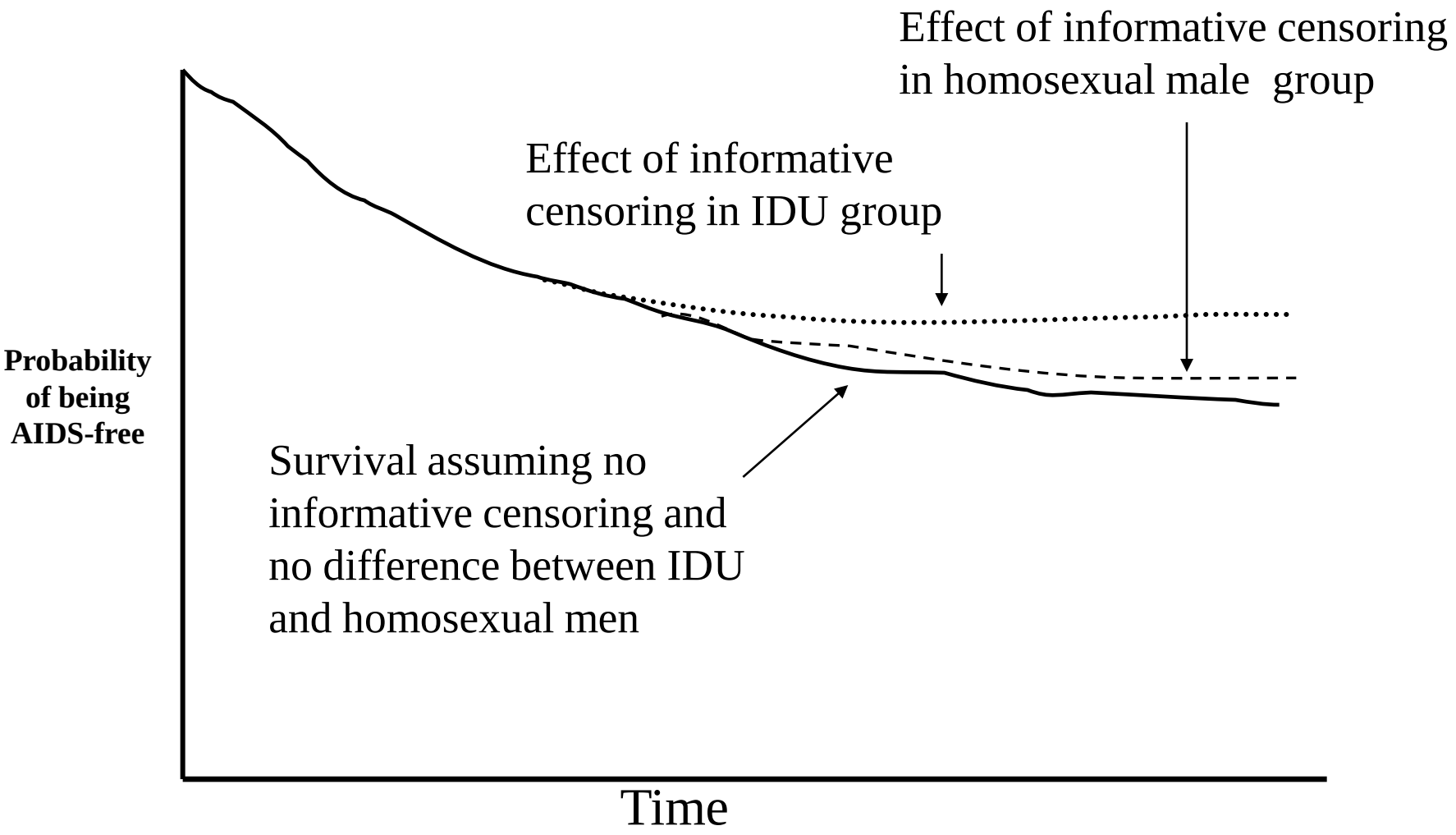
- Other examples:
  - Bias in using the general population as a comparison group for occupational cohorts
  - Bias due to differential drop-out rates among exposed and unexposed
    - E.g. cohort study on progression to AIDS
  - Bias when the analysis is restricted to individuals with complete follow-up
    - E.g. cohort studies on smoking and dementia

Selection Bias: Cohort Studies

Example: Cohort study of progression to AIDS: IV drug users (IDU) vs homosexual men

- In general, getting sicker is a common reason for loss to follow-up
- Therefore, persons who are lost to follow-up have different AIDS incidence than those who remain (i.e., informative censoring)
- In general, IDU more likely to be lost to follow-up - at any given level of feeling sick
- Therefore, the degree of informative censoring differs across exposure groups (IDU vs homosexual men)
- Results in selection bias: underestimates the incidence of AIDS in IDU relative to homosexual men

Selection Bias: Cohort Studies



IDU who are sick are likely to be lost during follow-up
Those who remain are likely to have a better prognosis

Selection Bias: Cohort Studies

Example: Cohort studies of smoking and dementia:

Cigarette Smoking and Dementia

Potential Selection Bias in the Elderly

Miguel A. Hernán,[a] *Alvaro Alonso,*[b] *and Giancarlo Logroscino*[a]

Abstract: We conducted a systematic review of published prospective studies that estimated the association between smoking and the incidence of Alzheimer disease and dementia. The relative rate for smokers versus nonsmokers ranged from 0.27 to 2.72 for Alzheimer disease (12 studies) and from 0.38 to 1.42 for dementia (6 studies). The minimum age at entry (range: 55–75 years) explained much of the between-study heterogeneity in relative rates. We conjecture that selection bias due to censoring by death may be the main explanation for the reversal of the relative rate with increasing age.

(*Epidemiology* 2008;19: 448–450)

prospective OR epidemiolog*)." We excluded studies that relied exclusively on death certificates to ascertain the dementia diagnosis.[4,5]

Table 1 summarizes the characteristics of the 12 studies that met our criteria.[6–17] The relative rate (RR) for smokers versus nonsmokers ranged from 0.27 to 2.72 for Alzheimer disease (12 studies) and from 0.38 to 1.42 for dementia (6 studies). We hypothesized that part of this between-study heterogeneity could be explained by the between-study differences in minimum age at entry (range: 55–75 years).

Figure 1 plots the log RR of Alzheimer disease versus

Hernan et al. Epidemiology 2008

Selection Bias: Cohort Studies

- Smoking harmful in studies that enrolled younger subjects, and appeared protective in studies that enrolled the oldest subjects
- Two possible explanations:
  - First, the effect of cigarette smoking on the risk of dementia is modified by age: smoking harmful at younger ages, beneficial at older ages.
  - Second, the effect of cigarette smoking is harmful overall but appears beneficial at older ages because of selection bias, eg, most smokers who are susceptible to developing dementia due to their smoking do so by age 75, and thus the group of 75-year-olds without dementia at baseline is depleted of susceptible smokers.

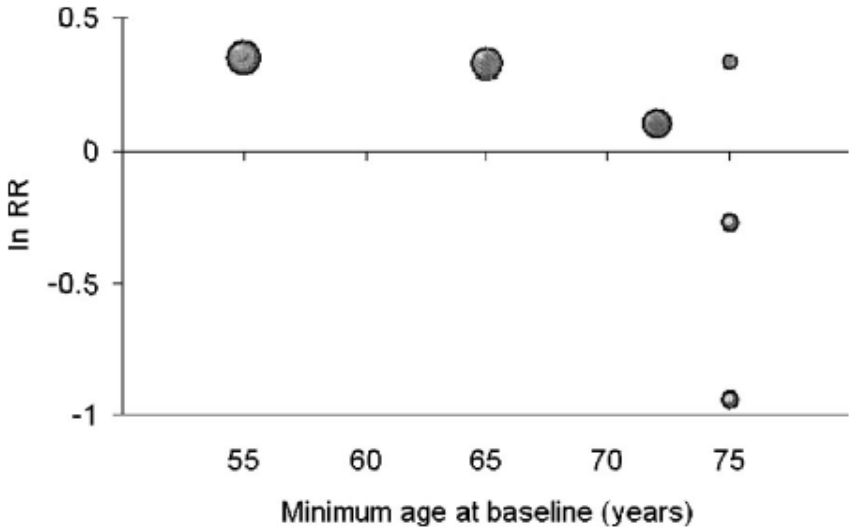


FIGURE 2. Log RR of dementia by the minimum age at baseline in the study. The area of the circle is proportional to the precision (1/variance) of the log RR estimate.

Hernan et al. Epidemiology 2008

Selection bias in case-control studies

- Sources:
  - Bias in selection of cases
    - Cases are not derived from a well defined study base (or source population)
  - Bias in selection of controls
    - Controls should provide an unbiased sample of the exposure distribution in the study base
    - Control selection is a more important issue than case selection!

Selection bias in case-control studies

- Examples:
  - Bias due to control selection:
    - Case-control study tampons and toxic shock syndrome (Reingold AL et al. Rev Infect Dis. 1989 Jan-Feb;11 Suppl 1:S35-41)
    - Case-control study on coffee drinking and pancreatic cancer (MacMahon et al. N Engl J Med. 1981 Mar 12;304(11):630-3)
    - Bias due to selection of hospital controls

Selection bias in case-control studies

- **Risk factors for menstrual toxic shock syndrome: results of a multistate case-control study.**

  For assessment of current risk factors for developing toxic shock syndrome (TSS) during menstruation, a case-control study was performed
- Cases with onset between 1 January 1986 and 30 June 1987 were ascertained in six study areas with active surveillance for TSS
- Age-matched controls were selected from among each patient's friends and women with the same telephone exchange
- Of 118 eligible patients, 108 were enrolled, as were 185 "friend controls" and 187 telephone exchange-matched controls

Reingold AL et al. Rev Infect Dis. 1989 Jan-Feb;11 Suppl 1:S35-41

Selection bias in case-control studies

- Risk factors for menstrual toxic shock syndrome: results of a multistate case-control study

- **Results for tampon use as a risk factor:**
  - OR when both control groups were combined = 29
  - OR when friend controls were used = 19
  - OR when neighborhood controls were used = 48

- Why did use of friend controls produce a lower OR?
  - Friend controls were more likely to have used tampons than were neighborhood controls (71% vs. 60%)

Reingold AL et al. Rev Infect Dis. 1989 Jan-Feb;11 Suppl 1:S35-41

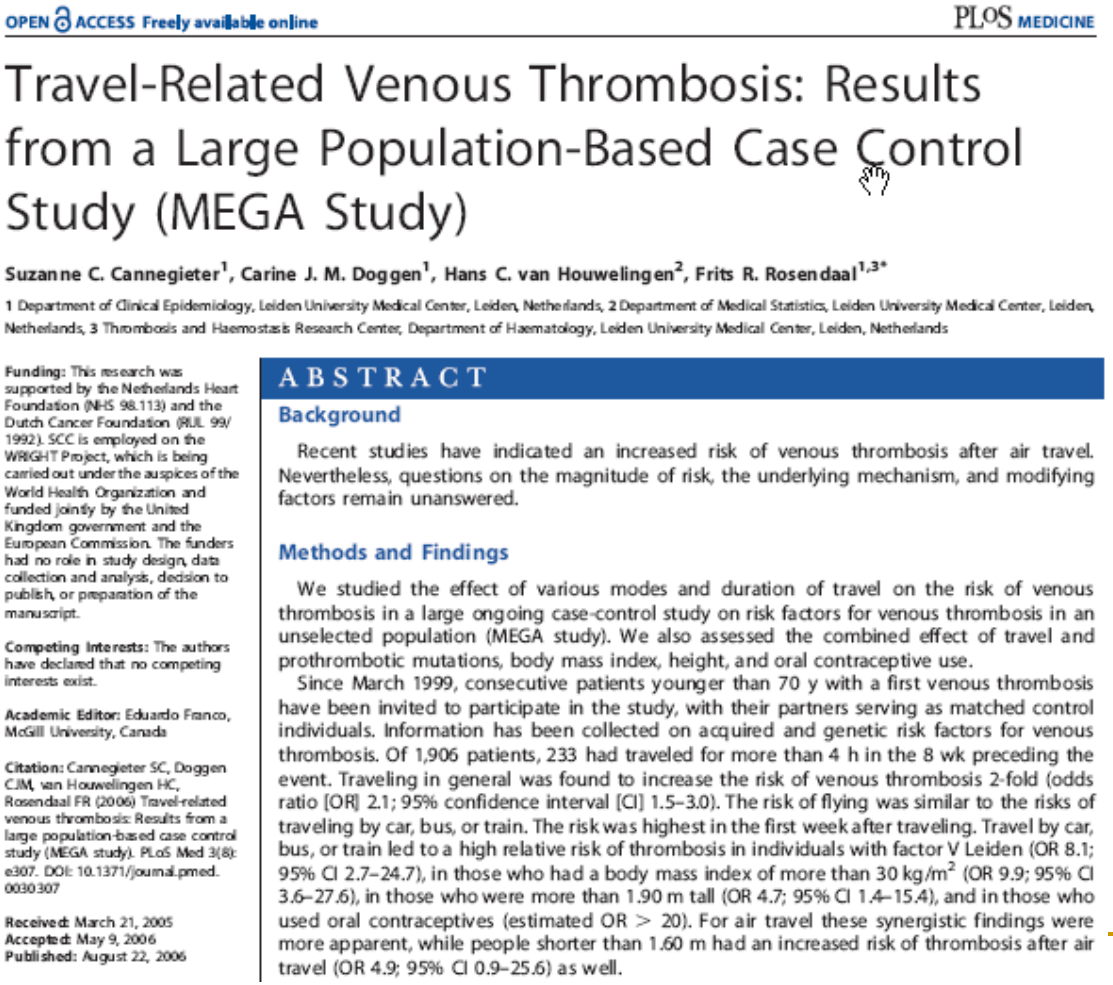
Direction of bias

		Case		Control
Exposure	Yes	a	⟺	b
	No	c		d

OR = ad / bc

If cases and controls share similar exposures (e.g. friend controls), then a and b will tend to be nearly the same -- this will bias the OR towards 1 (towards null)

In general, use of partners/spouses or friends as controls can result in bias



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.

In this case-control study, partners were controls, but couples often travel together and could have similar travel exposures!

Selection bias in case-control studies

630 THE NEW ENGLAND JOURNAL OF MEDICINE March 12, 1981

COFFEE AND CANCER OF THE PANCREAS

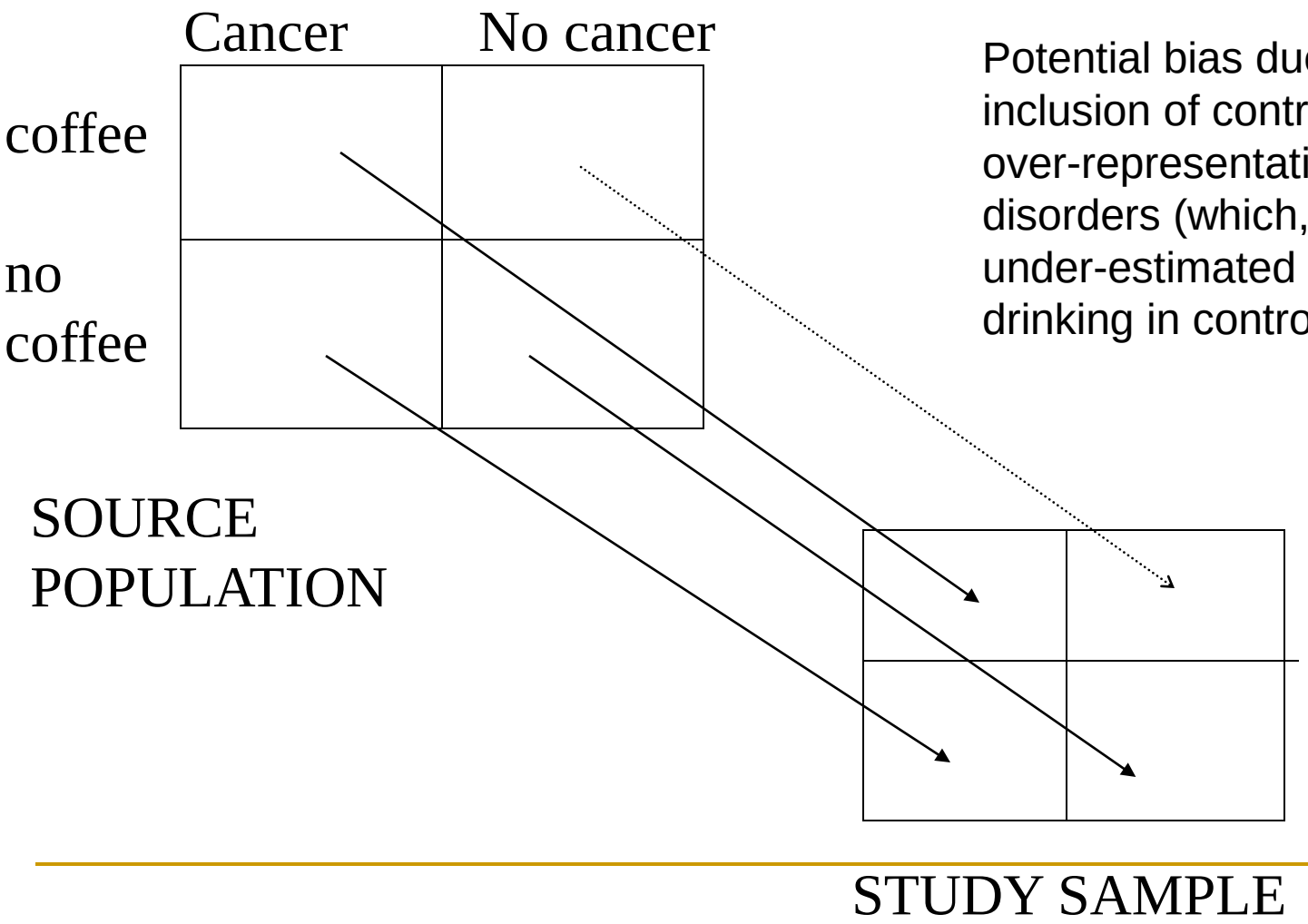
BRIAN MACMAHON, M.D., STELLA YEN, M.D., DIMITRIOS TRICHOPOULOS, M.D., KENNETH WARREN, M.D., AND GEORGE NARDI, M.D.

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 in this study were selected from a group of patients hospitalized by the same physicians who had diagnosed and hospitalized the cases' disease. The idea was to make the selection process of cases and controls similar. It was also logistically easier to get controls using this method. However, as the exposure factor was coffee drinking, it turned out that patients seen by the physicians who diagnosed pancreatic cancer often had gastrointestinal disorders and were thus advised not to drink coffee (or had chosen to reduce coffee drinking by themselves). So, this led to the selection of controls with higher prevalence of gastrointestinal disorders, and these controls had an unusually low odds of exposure (coffee intake). These in turn may have led to a spurious positive association between coffee intake and pancreatic cancer that could not be subsequently confirmed.

MacMahon et al. N Engl J Med. 1981 Mar 12;304(11):630-3

Case-control Study of Coffee and Pancreatic Cancer: Selection Bias



Potential bias due to inclusion of controls with over-representation of GI disorders (which, in turn, under-estimated coffee drinking in controls)

Direction of bias

		Case	Control
Exposure	Yes	a	b
	No	c	d

OR = ad / bc

If controls have an unusually low prevalence of exposure, then b will tend to be small -- this will bias the OR away from 1 (over-estimate the OR)

Coffee and cancer of the pancreas: Use of population-based controls

- Gold et al. *Cancer* 1985

	Case	Control
Coffee: ≥ 1 cup day	84	82
No coffee	10	14

OR= (84/10) / (82/14) = 1.4 (95% CI, 0.55 - 3.8)

So, when population-based controls were used, there was no strong association between coffee and pancreatic cancer

For a more in-depth analysis of this case study, see B-File #2

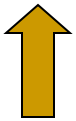
Bias File 2. Should we stop drinking coffee? The story of coffee and pancreatic cancer

Bias due to selection of hospital controls

- Example:
- In a case-control study of smoking and chronic obstructive pulmonary disease (COPD), controls were selected from the same hospital with other lung diseases (e.g. tuberculosis, lung cancer, occupational lung diseases).
- The authors found a weak association between smoking and COPD
- What is the problem with this study??
  - Smoking causes many diseases resulting in higher hospitalization rate of smokers
  - Hospital controls do not represent the prevalence of exposure (smoking) in the source population from which cases of COPD arose
  - Also, hospitalized people tend to have multiple diseases, and this can result in the distortion of the exposure frequencies in hospitalized controls (Berkson's bias)

Direction of bias due to hospitalized controls

		Case	Control
Exposure	Yes	a	b
	No	c	d



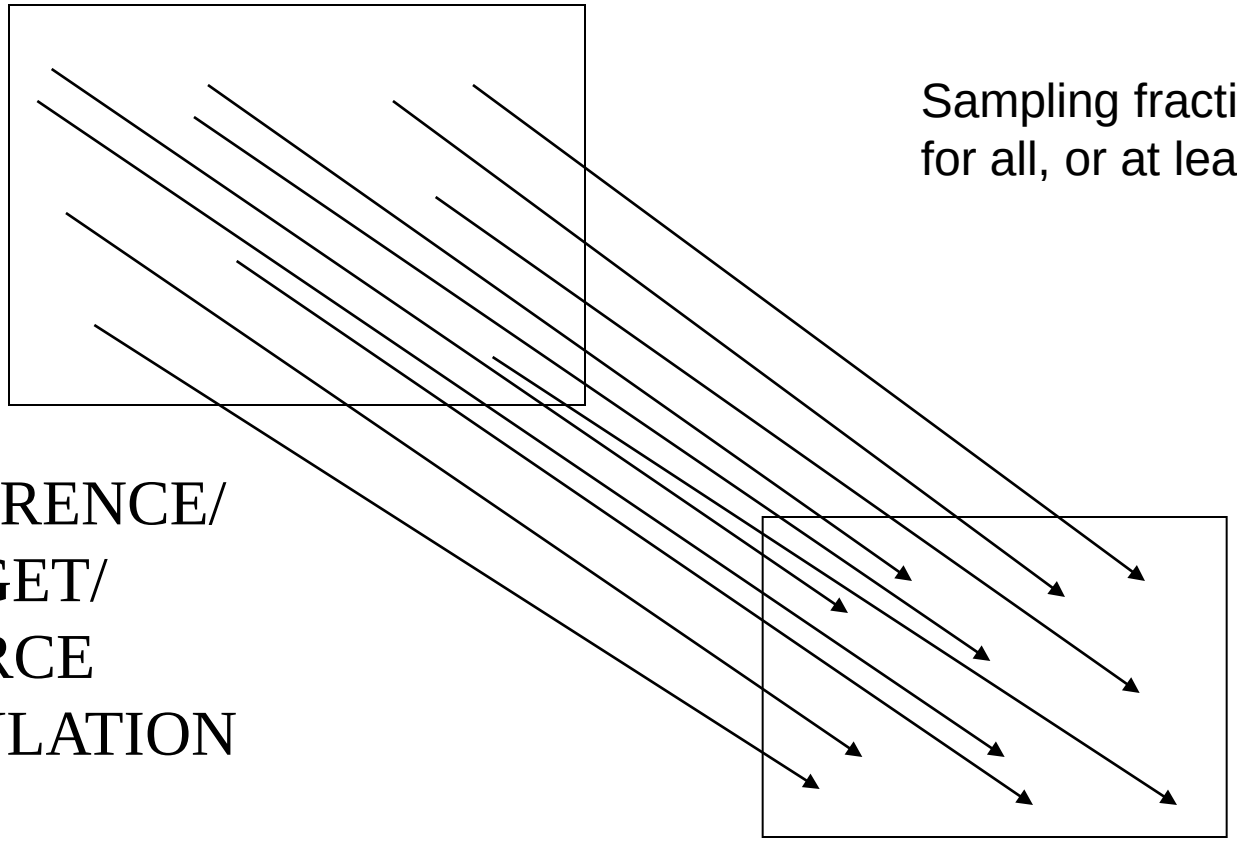
OR = ad / bc

If controls have an unusually high prevalence of exposure, then b will tend to be large -- this will bias the OR towards 1 (under-estimate the OR)

Selection bias in cross-sectional studies

- Sources:
  - Bias due to sampling
    - Selection of “survivors” or “prevalent” cases
    - Non-random sampling schemes
    - Volunteer bias
    - Membership bias
  - Bias due to non-participation
    - Non-response bias

Descriptive Study: Unbiased Sampling

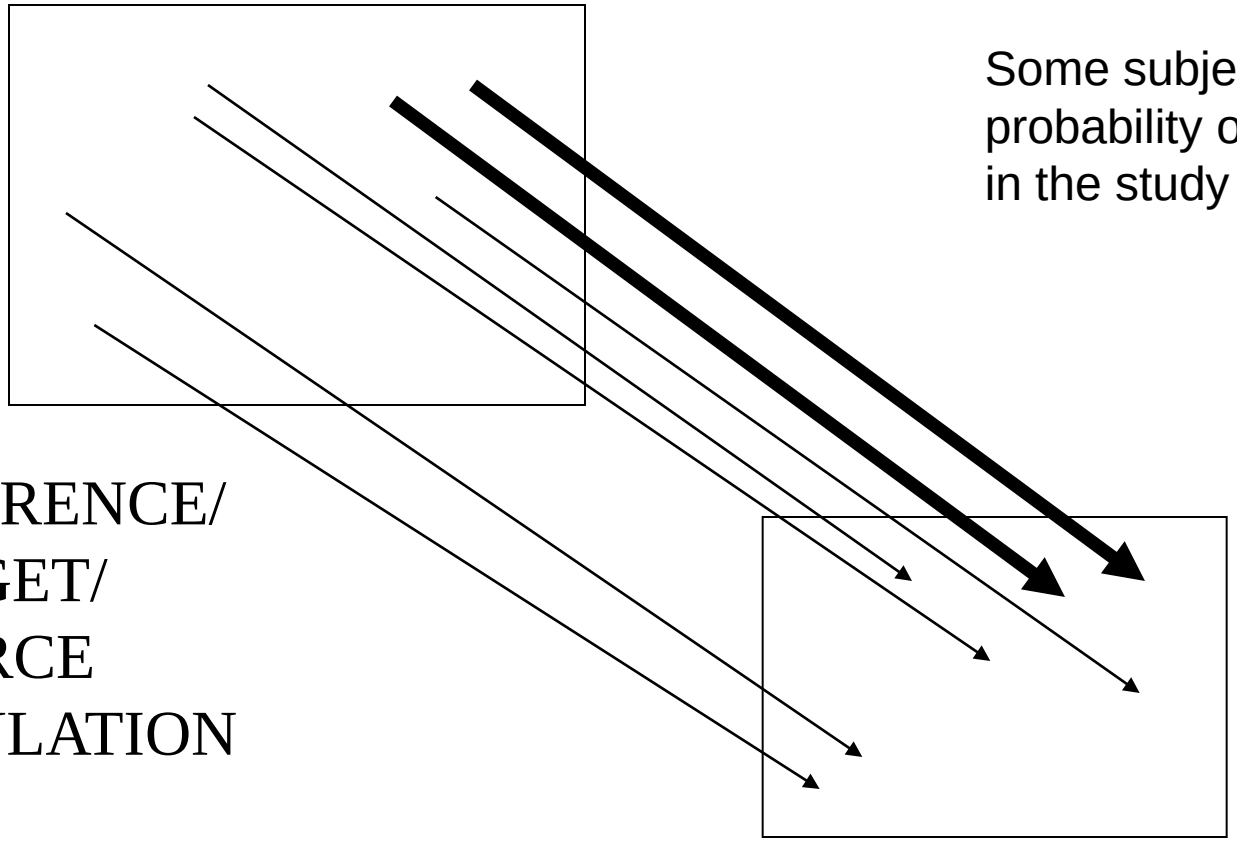


Sampling fraction is equal for all, or at least known

REFERENCE/
TARGET/
SOURCE
POPULATION

STUDY SAMPLE

Descriptive Study: Biased sampling

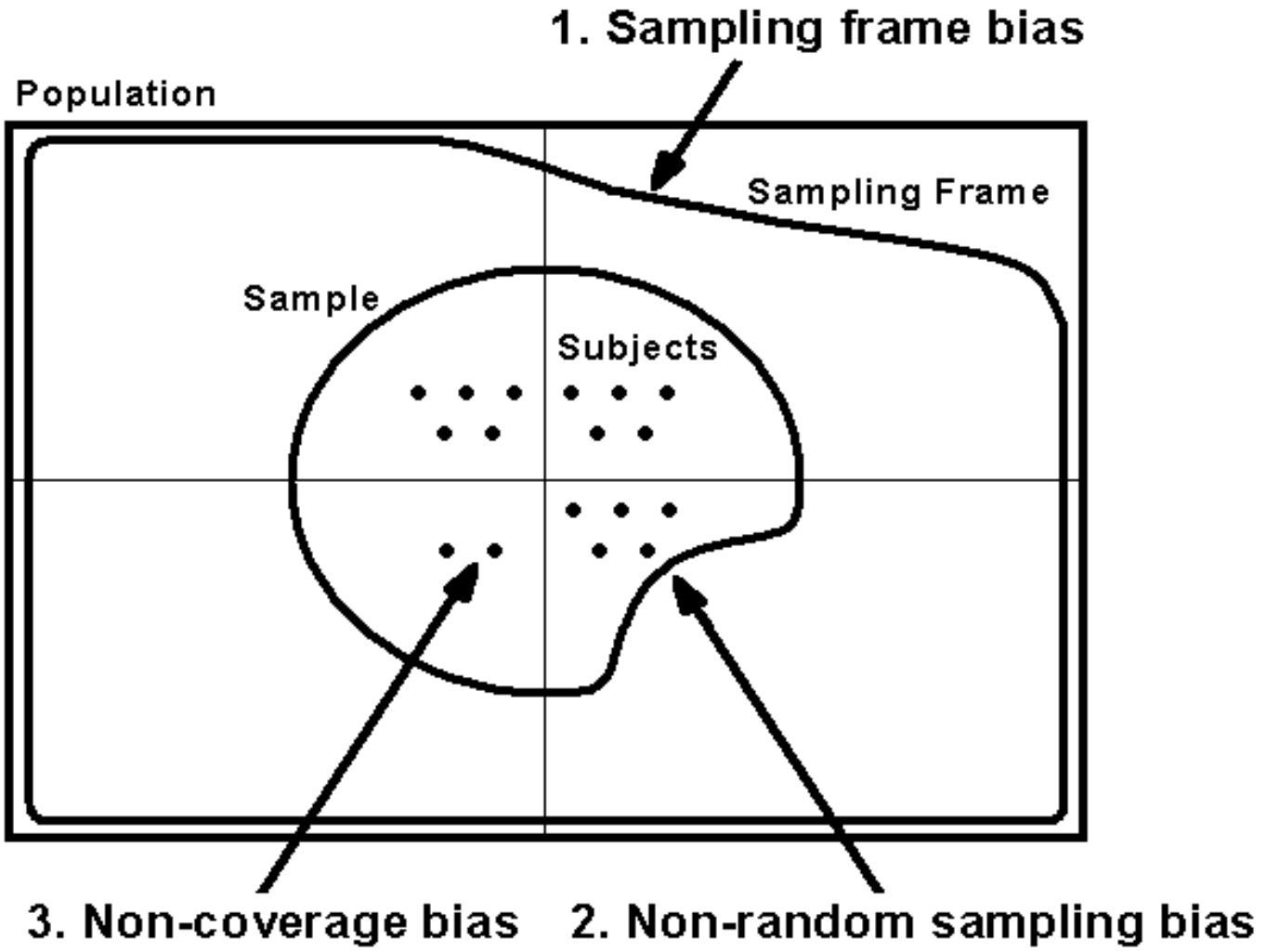


Some subjects have a higher probability of being included in the study sample

REFERENCE/
TARGET/
SOURCE
POPULATION

STUDY SAMPLE

Selection bias in sample surveys



Selection bias in telephone surveys

Table 1 Continuing and emerging challenges for telephone survey research

Ongoing challenges	New and emerging challenges
Selecting participants	
Sampling	Cell phone sampling
Telephone coverage	Number portability
Response rates	Answering machines
Participation rates	Caller ID
Call scheduling	Privacy managers and call blocking
Collecting information	
Reliable and valid responses	Privacy and confidentiality
Mode effects	Respondent burden

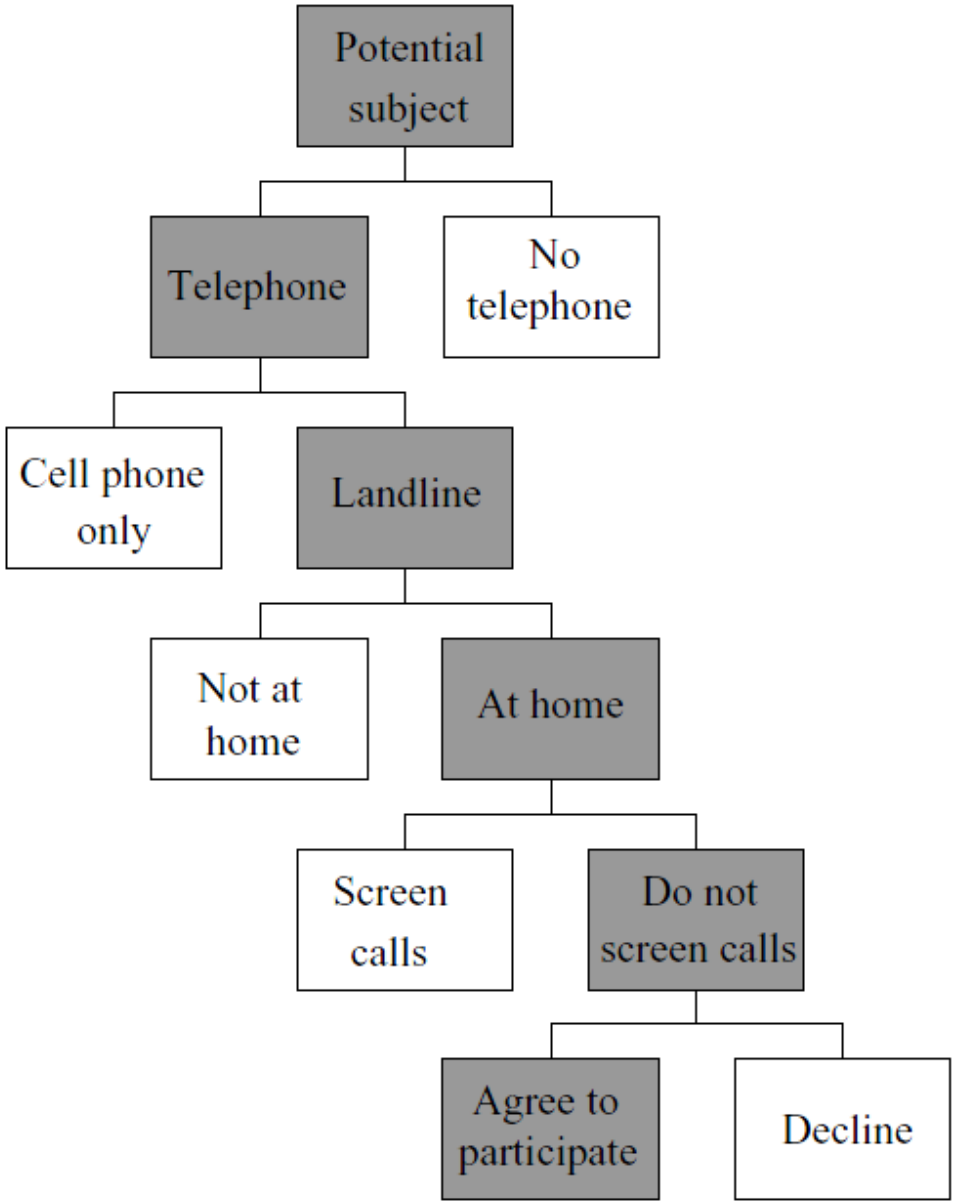


Figure 1

Steps in the selection of participants in telephone surveys.

Annu.Rev.PublicHealth 2007.28:113–26

Selection bias in cross-sectional studies

- Examples:
  - Bias due to sampling:
    - healthy worker effect (or bias): survey on occupational lung disease (silicosis among stone quarry workers)
    - Volunteer bias: bias in screening programs (e.g. leukemia among nuclear test observers)
  - Non-response bias
    - Survey on prevalence of self-reported diabetes (Pai et al. 1999)
  - Survivor bias
    - Study to determine neurological status of patients who had survived after CPR in a hospital in India (Rajagopalan et al, 1999)

Example: Study on mental health disorders among marines deployed to combat

- Research studies have identified heightened psychiatric problems among veterans of Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF).
- Study done to determine incidence rates of diagnosed mental disorders in a cohort of Marines deployed to combat during OIF or OEF in 2001–2005 and to compare these with mental disorder rates in two historical and two contemporary military control groups.
- All psychiatric conditions except post-traumatic stress disorder occurred at a lower rate in combat-deployed personnel than in personnel who were not deployed to a combat zone.

Larson et al. AJE 2008

‘Healthy Warrior Effect’ [belongs to the same family as ‘Healthy Worker Bias’

Psychiatric Diagnoses in Historic and Contemporary Military Cohorts: Combat Deployment and the Healthy Warrior Effect

Gerald E. Larson[1], Robyn M. Highfill-McRoy[1,2], and Stephanie Booth-Kewley[1]

[1] Behavioral Science and Epidemiology Department, Naval Health Research Center, San Diego, CA.
[2] Science Applications International Corporation, San Diego, CA.

Received for publication August 17, 2007; accepted for publication January 10, 2008.

Research studies have identified heightened psychiatric problems among veterans of Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF). However, these studies have not compared incidence rates of psychiatric disorders across robust cohorts, nor have they documented psychiatric problems prior to combat exposure. The authors' objectives in this study were to determine incidence rates of diagnosed mental disorders in a cohort of Marines deployed to combat during OIF or OEF in 2001–2005 and to compare these with mental disorder rates in two historical and two contemporary military control groups. After exclusion of persons who had been deployed to a combat zone with a preexisting psychiatric diagnosis, the cumulative rate of post-OIF/-OEF mental disorders was 6.4%. All psychiatric conditions except post-traumatic stress disorder occurred at a lower rate in combat-deployed personnel than in personnel who were not deployed to a combat zone. The findings suggest that psychiatric disorders in Marines are diagnosed most frequently during the initial months of recruit training rather than after combat deployment. The disproportionate loss of psychologically unfit personnel early in training creates a "healthy warrior effect," because only those persons who have proven their resilience during training remain eligible for combat.

Larson et al. AJE 2008

Bias due to non-response

- Survey to estimate prevalence of self-reported chronic diseases in a city in India (Pai et al, 1999)
- 705 adults were interviewed (of an eligible population of 808)
  - 29.1% had been told (by a doctor or health professional) that they had hypertension
- Proxy data was obtained for 32 of the non-responders [who could never be contacted, despite repeated attempts]
  - 45.8% of non-responders had self reported hypertension
  - If these people had been included, the overall prevalence would have been higher

Prevalence of self-reported hypertension	Responders n=705	Non-responders n=32
	29.1%	45.8%

Leukemia Incidence Among Observers of a Nuclear Bomb Test (Volunteer bias)

Caldwell GG et al. *JAMA* 1980

- Smokey Atomic Test in Nevada
- 76% of troops at site was later found; occurrence of leukemia determined

82% contacted by the investigators

18% contacted the investigators on their own

4.4 greater risk of leukemia than those contacted by the investigators

More on selection probabilities

Suppose in a study of asbestos exposure and lung cancer the exposure is distributed among the cases and controls in the target population as follows:

	Diseased	Nondiseased
Exposed	100	200
Unexposed	100	400

The true OR in the target population is (100 x 400) / (100 x 200) = 2.0.

More on selection probabilities

If the selection probabilities for all the cells in the table are equal at 90%, the 2x2 table of selection probabilities would look like the following.

	Diseased	Nondiseased
Exposed	• =90%	• =90%
Unexposed	• =90%	• =90%

cross product of α, β, γ, δ = 1

This cross-product is called Selection bias factor

And the 2x2 table of individuals in the case-control study will look like the following.

	Diseased	Nondiseased
Exposed	100 x .90=90	200 x .90=180
Unexposed	100 x .90=90	400 x .90=360

OR = (90 x 360) / (90 x 180) = 2.0

Is there selection bias?

More on selection probabilities

If the selection probabilities are unequal, but still proportional (that is, alpha / beta = gamma / delta), we still do not observe any selection bias in our study. If the selection probability is 90% among the diseased individuals and the selection probability is 70% among the nondiseased individuals the resulting 2x2 table would look like the following.

cross product of α, β, γ, δ = 1

	Diseased	Nondiseased
Exposed	100 x .90=90	200 x .70=140
Unexposed	100 x .90=90	400 x .70=280

OR = (90 x 280) / (90 x 140) = 2.0

Within cases and controls, the exposure odds is maintained

Is there selection bias?

More on selection probabilities

If however the selection probabilities are unequal, and also nonproportional, then selection bias will occur. The following table shows how selection bias occurs when the selection probability for the unexposed controls is different than that of the other three groups of study members.

cross product of $\alpha, \beta, \gamma, \delta$ =/= 1

	Diseased	Nondiseased
Exposed	100 x .90=90	200 x .90=180
Unexposed	100 x .90=90	400 x .70=280

Control group has higher odds of exposure (180/280) than the study Base (200/400)

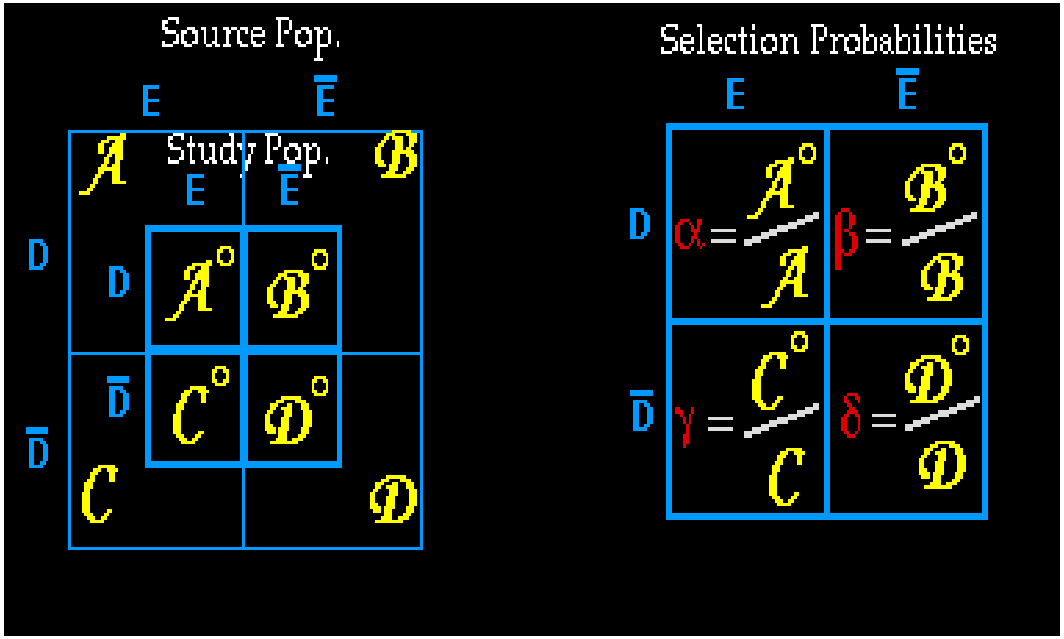
OR = (90 x 280) / (90 x 180) = 1.6

Within cases, the exposure odds is maintained
Within controls, the exposure odds is distorted

Is there selection bias?

Can selection bias be “fixed”?

- Not easy
  - Best avoided at the design stage; can try hard to retain participants in the study
- Can collect data to ‘estimate’ magnitude/direction of selection bias and do sensitivity analysis
  - e.g., collect data from a sample of non-respondents, and use this to do sensitivity analysis
- Effect estimates can be ‘adjusted’ if selection probabilities are known
  - Good sources: Kleinbaum’s ActivEpi book/CD & new book on bias analysis by Lash et al.



To adjust, we need selection probabilities.
But how do we get them??

Software programs for bias analysis (sensitivity analysis)

The Stata Journal (2008)
8, Number 1, pp. 29–48

A tool for deterministic and probabilistic sensitivity analysis of epidemiologic studies

Nicola Orsini
Division of Nutritional Epidemiology
Institute of Environmental Medicine
Karolinska Institutet
Stockholm, Sweden
nicola.orsini@ki.se

Rino Bellocco
Department of Statistics
University of Milano-Bicocca
Milano, Italy

Matteo Bottai
Department of Epidemiology and Biostatistics
Arnold School of Public Health
University of South Carolina
Columbia, SC

Alicja Wolk
Division of Nutritional Epidemiology
Institute of Environmental Medicine
Karolinska Institutet
Stockholm, Sweden

Sander Greenland
Departments of Epidemiology and Statistics
University of California, Los Angeles
Los Angeles, CA

Book on bias analysis (sensitivity analysis)

Applying Quantitative Bias Analysis to Epidemiologic Data
Springer, 2009
Lash, Timothy L., **Fox**, Matthew P., **Fink**, Aliza K.

Includes SAS codes for programs

Readings this week

- Rothman: Chapter 5: Biases in Study Design
- Gordis:
  - Chapter 14: From Association to Causation
  - Chapter 15: More on Causal Inferences: Bias, Confounding, and Interaction
- Article:
  - ERIC Notebook handout on Selection Bias, UNC

An Introduction to

QUANTUM Gradnamics

Although Quantum Gradnamics explains many of the phenomena in pursuing a Ph.D., most aspiring scientists still object to such an uncertain and probabilistic description of academic reality.

The Austrian scientist Erwin Schrödinger was particularly uncomfortable not knowing whether he would ever graduate or not, and illustrated this with his now famous thought experiment known as "Schrödinger's Cubicle."

According to the experiment, grad students exist in a state of both productivity and unproductivity (many students do report feeling like...

WWW.PHDCOMICS.COM
JORGE CHAM © 2007

"I had two wives at the same time. How's that for pressure to graduate!"
- Erwin Schrödinger

Schrödinger's Cubicle

1. Place grad student inside closed cubicle
2. Set up computer, coffee and internet connection
3. Wait a few years

...they're in limbo the whole time). Only direct intervention reveals whether or not an enormous amount of time has been wasted, a phenomenon known as "expectation collapse".

Einstein was also uncomfortable with this indeterminate view of academia and openly disagreed with the Copenhagen Interpretation, which states that graduation is an entirely random process. In deciding whether or not to graduate a student, Einstein famously said, "Professors don't throw dice (do they?)."

More recent theories describe grad students as soggy strings of ramen noodles, which is just as useful.

(thanks to Wikipedia for all the background info)

An Introduction to

QUANTUM Gradnamics

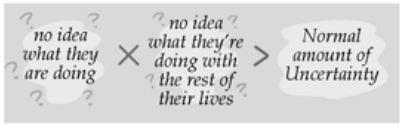
Another principal concept in Quantum Gradnamics is the observation that graduate students do not move toward graduation in a steady and continuous manner. Rather, they make progress through discrete bursts of random productivity called "wanta" (short for "want data") whose energy is proportional to the frequency of meetings with their advisor.

Grad students, or "p-ons" as Einstein called them, can only occupy a discrete number of energy states:

WWW.PHDCOMICS.COM
JORGE CHAM © 2007

"I discard all hope of predicting hitherto unpredictable quantities, such as my graduation."
- Werner Heisenberg

A direct consequence of this is the "Heisenberg Uncertainty Principle", perhaps the most well-known theorem of Quantum Gradnamics. Developed by Heisenberg during a particularly unproductive period in his graduate career, the principle states that it is not possible to know where a grad student is and where it is going at the same time:



When probed under pressure, a grad student will either blurt out what they are doing (but won't know if it means anything), or they will blurt out what they *plan* to do (but won't know *how* to do it). Simply put, there is an inherent degree of certainty and precision that is missing from their everyday life.

Heisenberg attributed this to the fact that meetings with professors are *non-communicative* (that is, the order in which orders are given doesn't tell you whether they are worth doing).