



# From Association to Causation:

# Deriving Inferences from Epidemiologic Studies

**Lina Wang**



Department of Epidemiology and Biostatistics  
School of Public Health  
Southeast University  
Phone: 025-83272569 (o)  
E-mail: [lnwang@seu.edu.cn](mailto:lnwang@seu.edu.cn)



# Approaches for Studying Disease Etiology

- Animal studies
  - In vitro systems
  - Randomized human trials
    - **LIMITATION:** cannot ethically expose people to suspected carcinogens
  - Non-randomized observational studies in human populations
    - e.g., cohort studies, case-control studies
- LIMITATION: May not be generalizable to whole human organisms**



# Natural (or “Unplanned”) Experiments

Investigators take advantage of opportunities to study groups of people who have been exposed for **non-study purposes**, such as **occupational cohorts** in specific industries or persons accidentally exposed to **toxic chemicals**.

- e.g., radiation exposure among residents of Hiroshima and Nagasaki, Japan, from atomic bombs in 1945



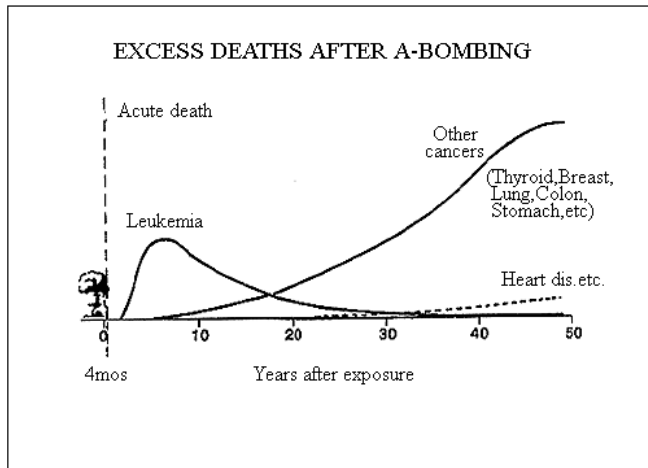
The mushroom cloud over Hiroshima after the dropping of Little Boy



The Fat Man mushroom cloud resulting from the nuclear explosion over Nagasaki rises 18 km (11 mi, 60,000 ft) into the air from the hypocenter.



**EXCESS DEATHS AFTER A-BOMBING**



Concerning the excess deaths due to atomic bomb radiation, those by leukemia were first noticed. Then, from about 10 years after the A-bombing up to now, the excess deaths due to other types of cancer have increased. In recent years, the excess deaths due to non-cancer diseases have also tended to show an increase. (Figure prepared by Radiation Effects Research Foundation.)



# Southeast University Percivall Pott Overview



alth

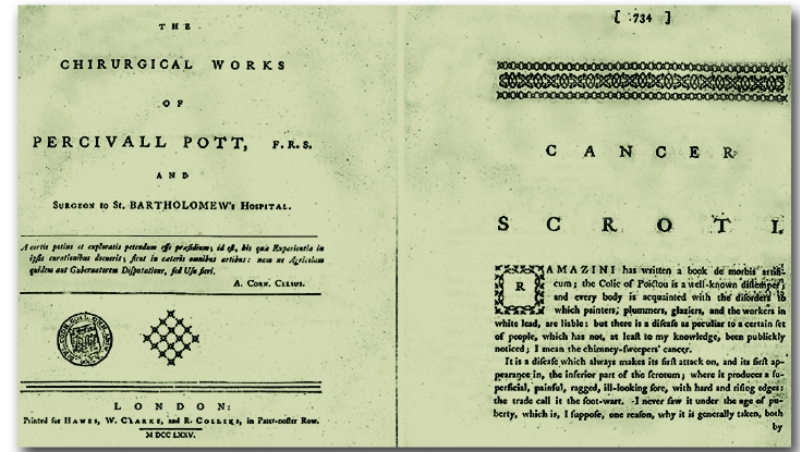


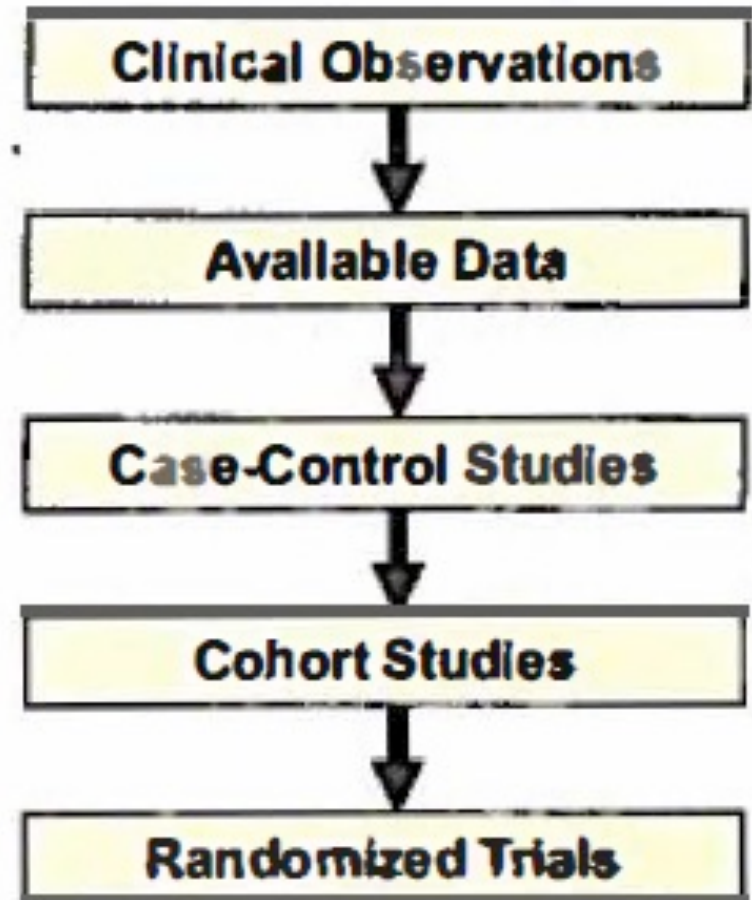
Figure 1. (left) Title page of Percivall Pott's highly regarded *Chirurgical Works* (Potts 1775). (right) First page of the chapter on scrotal cancers in chimney sweeps, generally held to be one of the first statements about chemical carcinogenesis and a very early formal reference to environmental medicine.





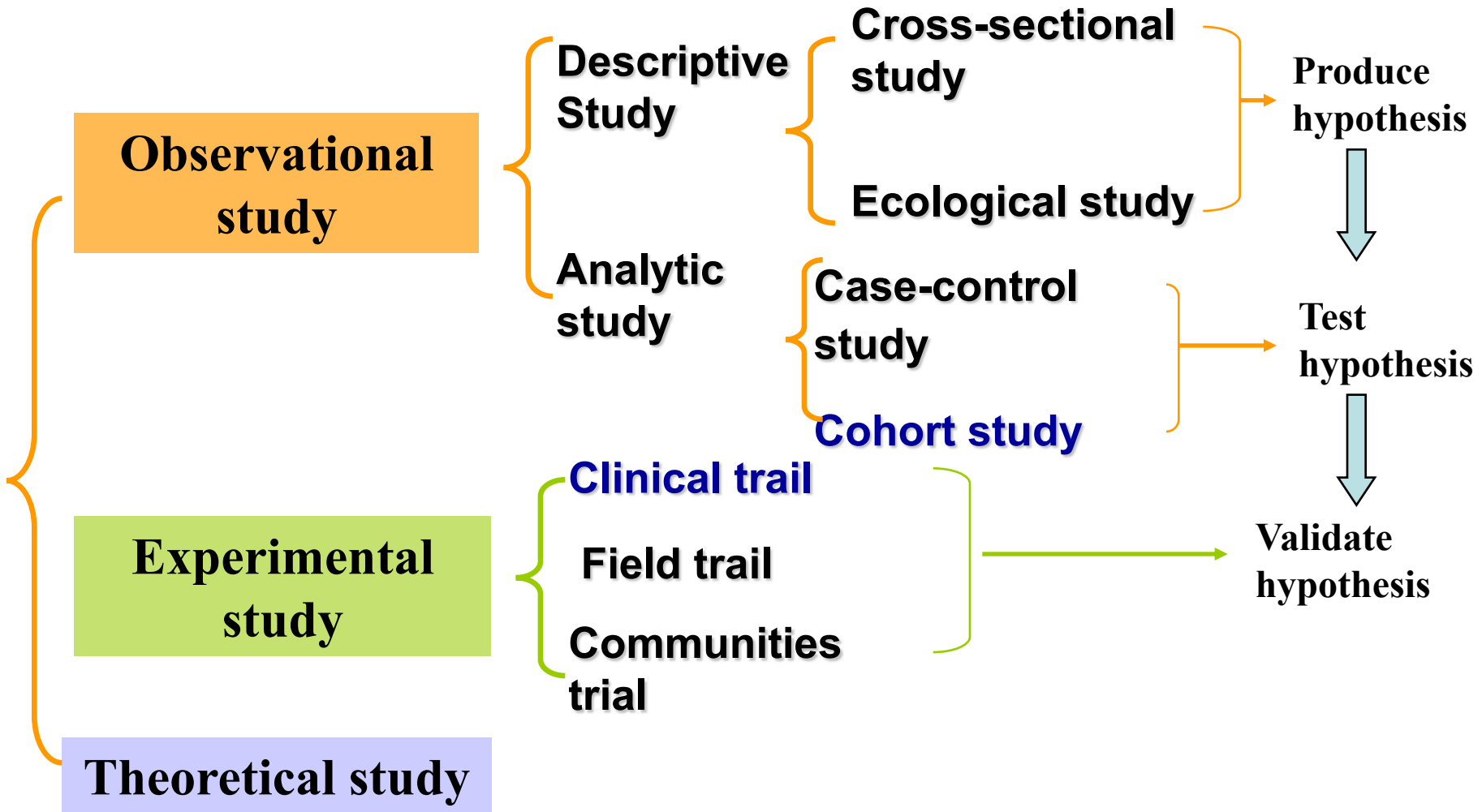
# Frequent Sequence of Studies in Human Populations

- 1) Determine whether there is an **association/correlation** between exposure/characteristic and risk of a disease.
- 2) If an association is demonstrated, determine whether the observed association is **causal**.





# Epidemiology Study Methods





# Concept Map

**Descriptive  
Epidemiology**



**Hypothesis**



**Analytical  
Epidemiology**

**RCT**

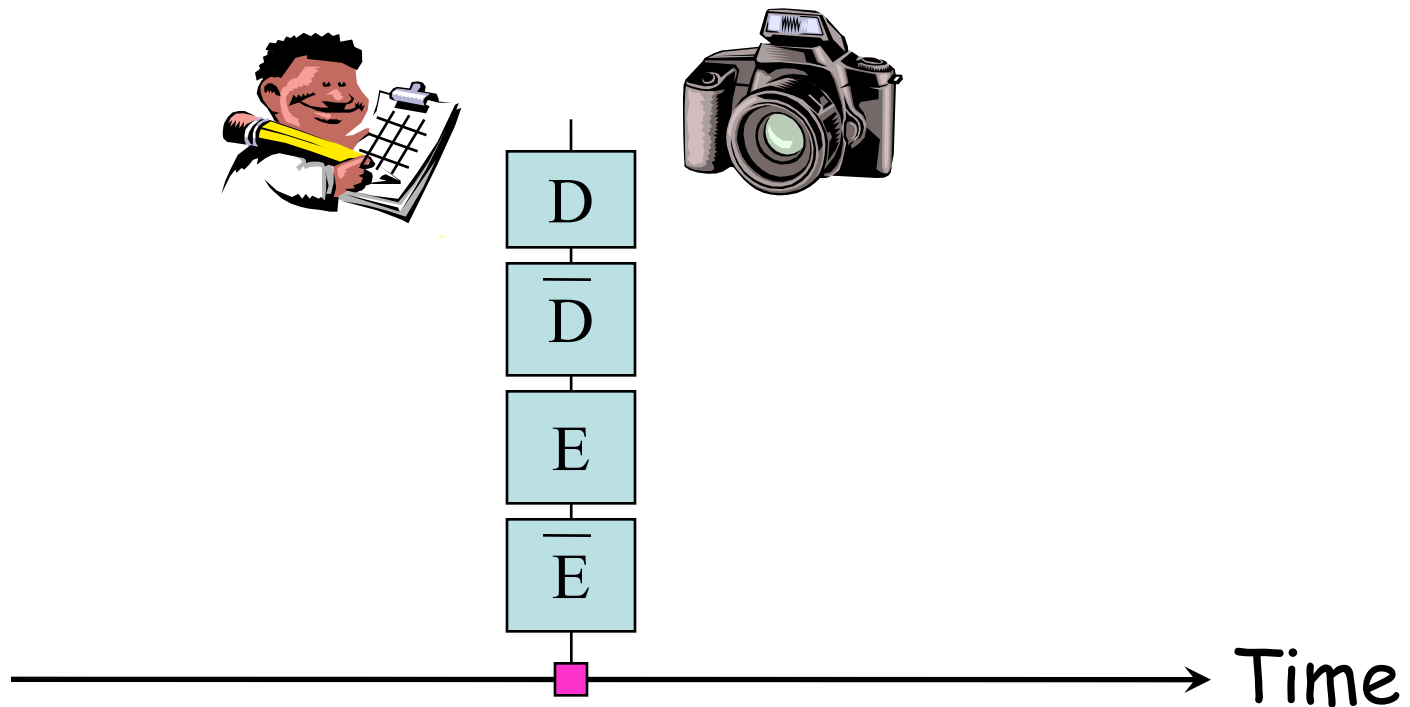


**Judging  
Causality**



**Prevention  
Strategies**

# cross-sectional study



**When Where Who**

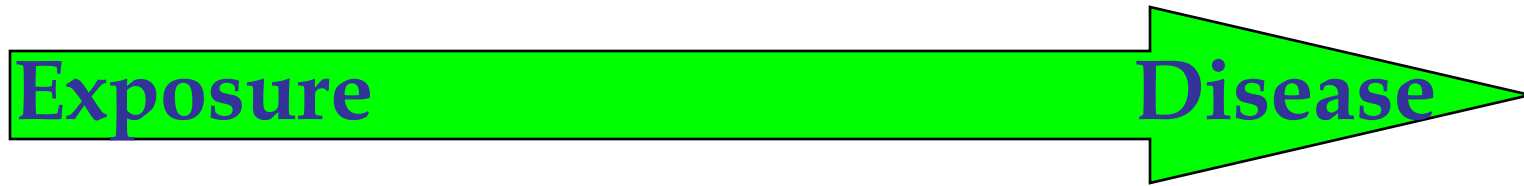


## Analytic Epidemiology

**The aspect of epidemiology concerned with the search for health-related causes and effects. Uses comparison groups, which provide baseline data, to quantify the association between exposures and outcomes, and test hypotheses about causal relationships.**



Time



*Cohort study*



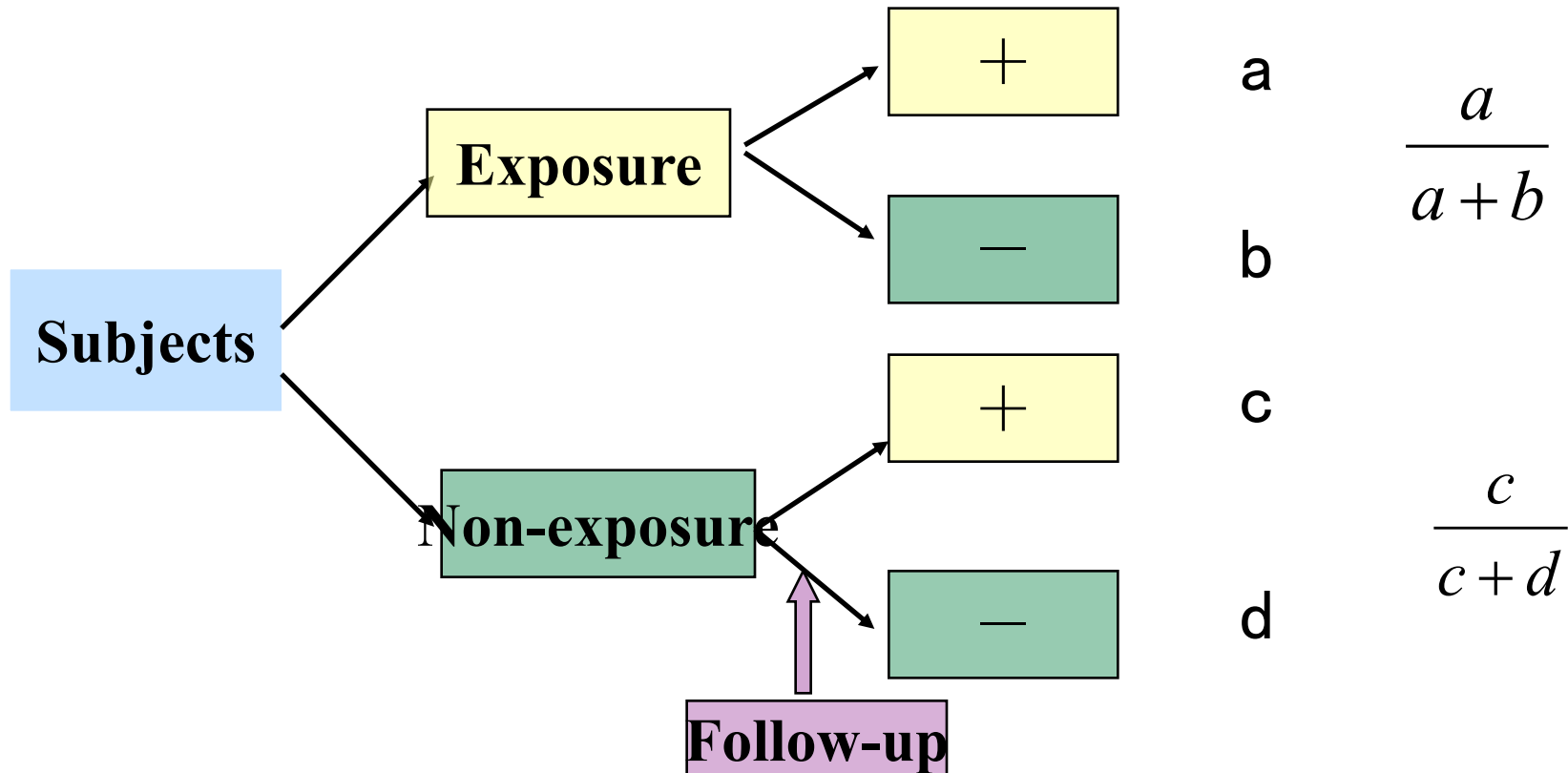
*Case control study*



# Cohort study

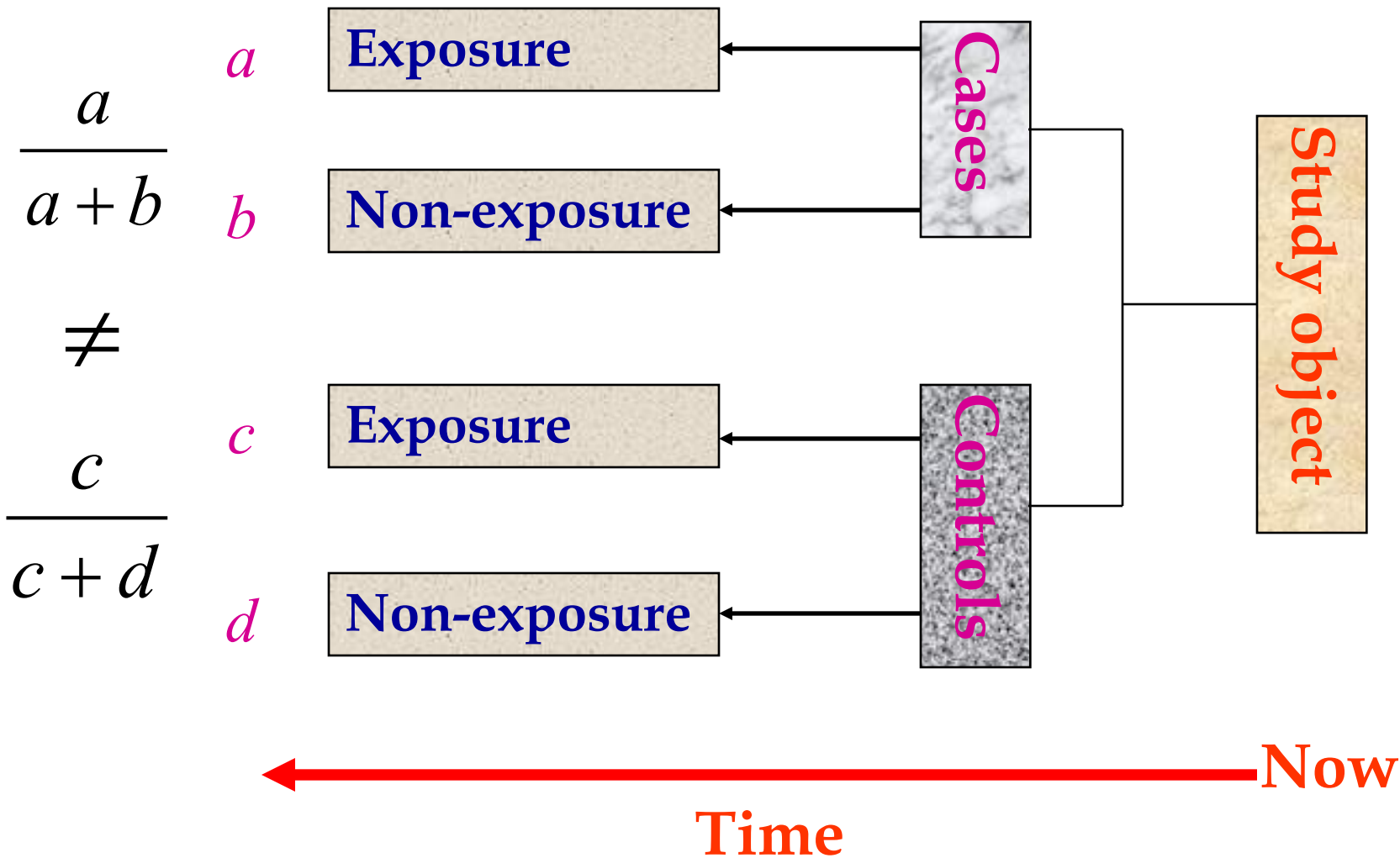
Now

Time →





# Case- control study

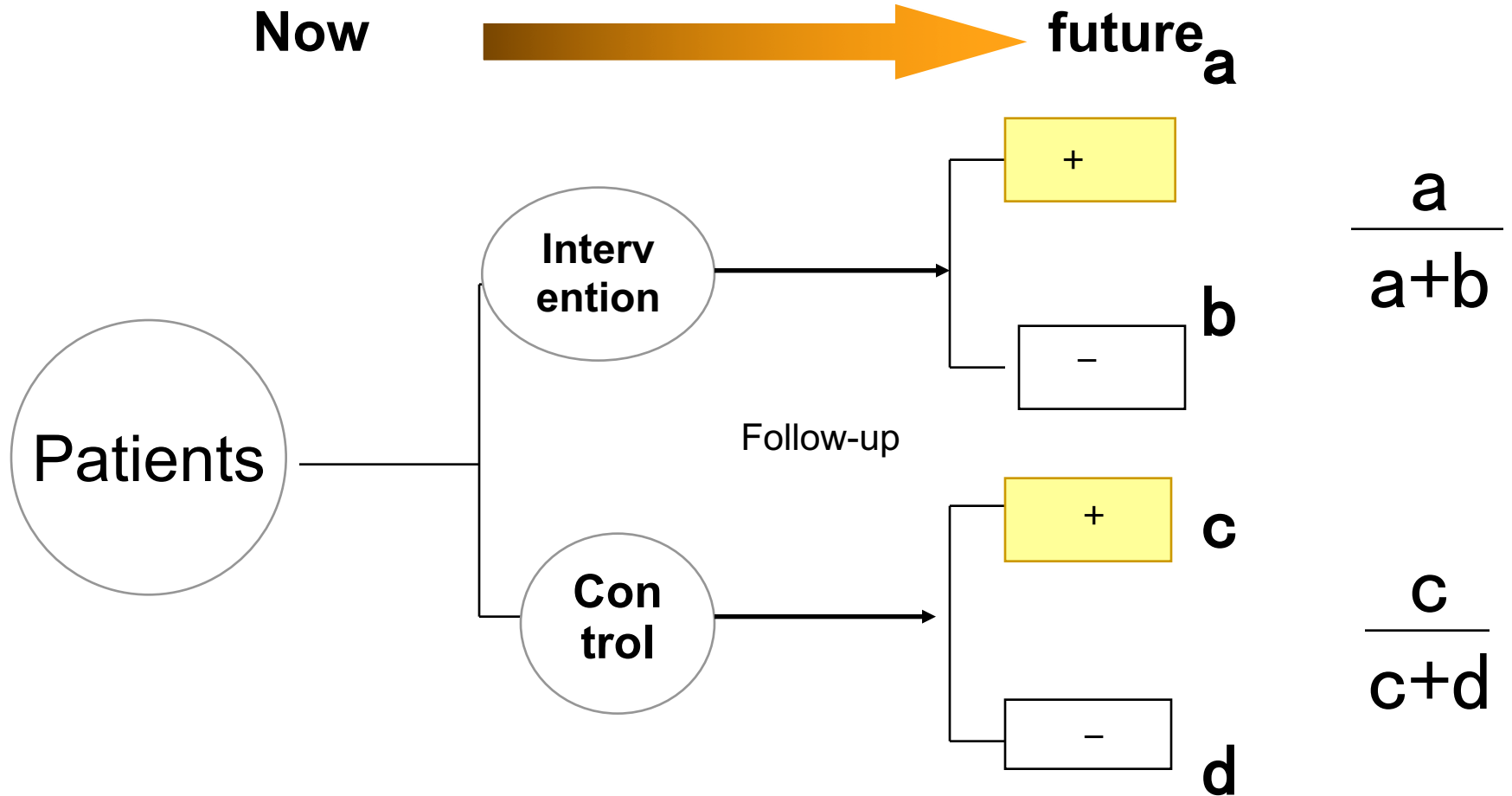




- A major limitation of **cross-sectional surveys** and **case-control studies** is difficulty in determining if exposure or risk factor preceded the disease or outcome.

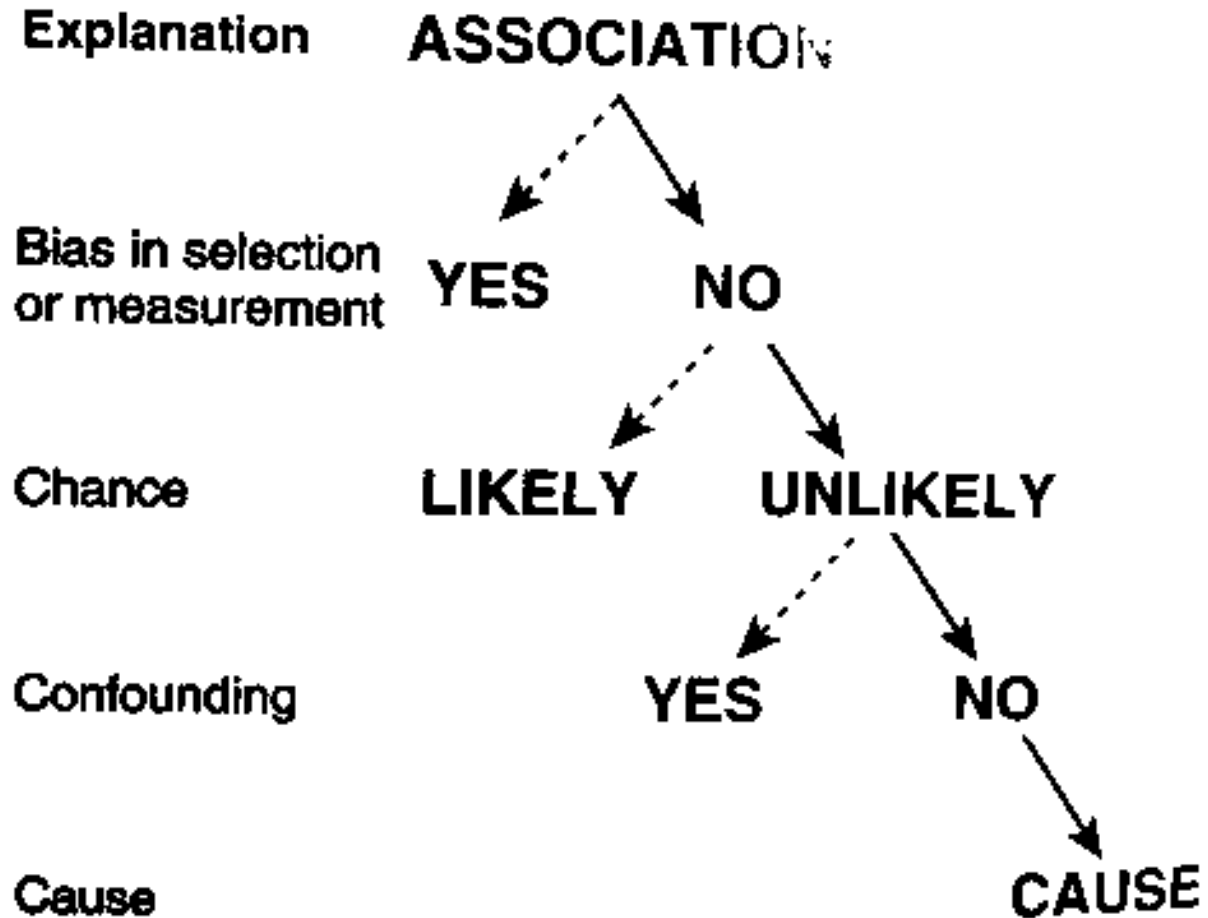


# Clinical trial



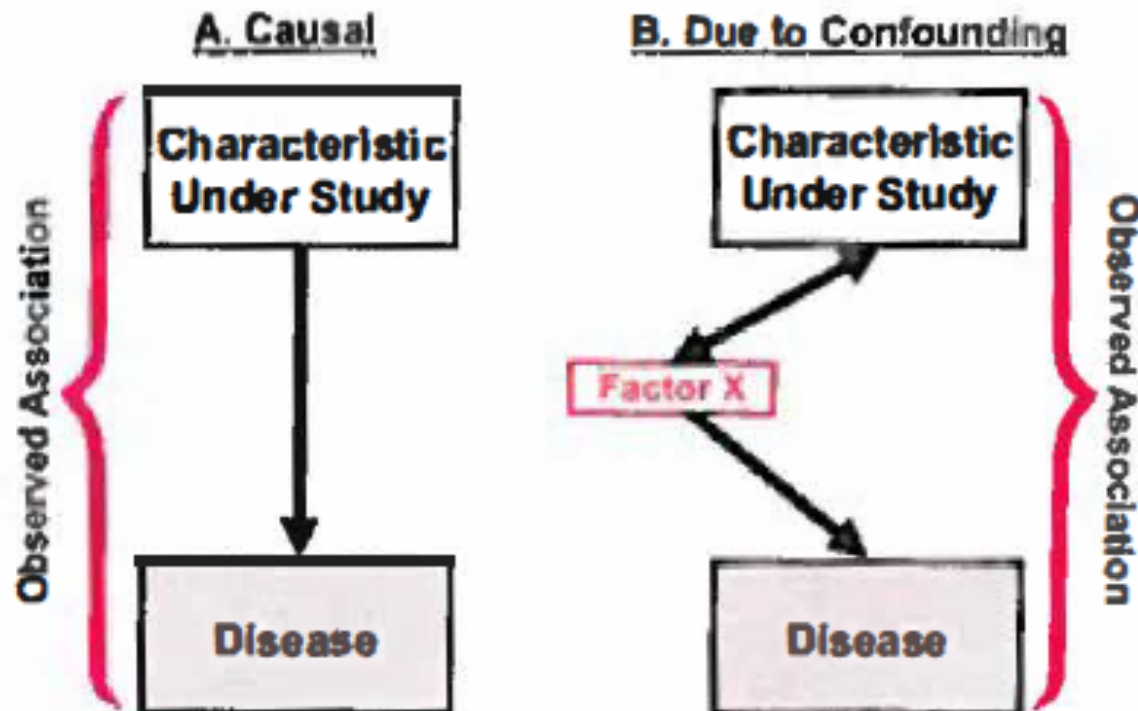


# Association and Cause



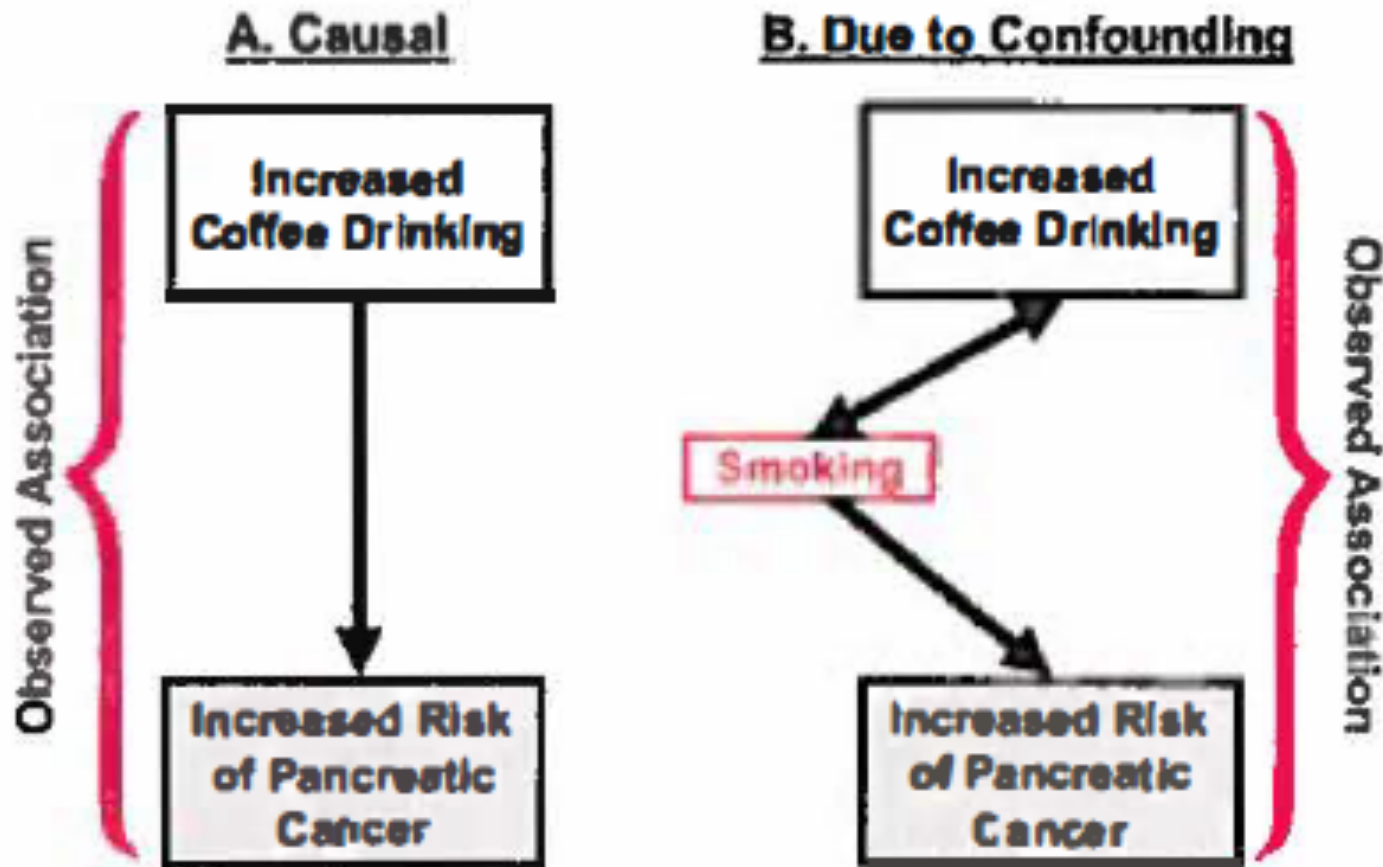
# Types of Associations

- **False** (spurious) associations
- **True** (real) associations—causal or noncausal



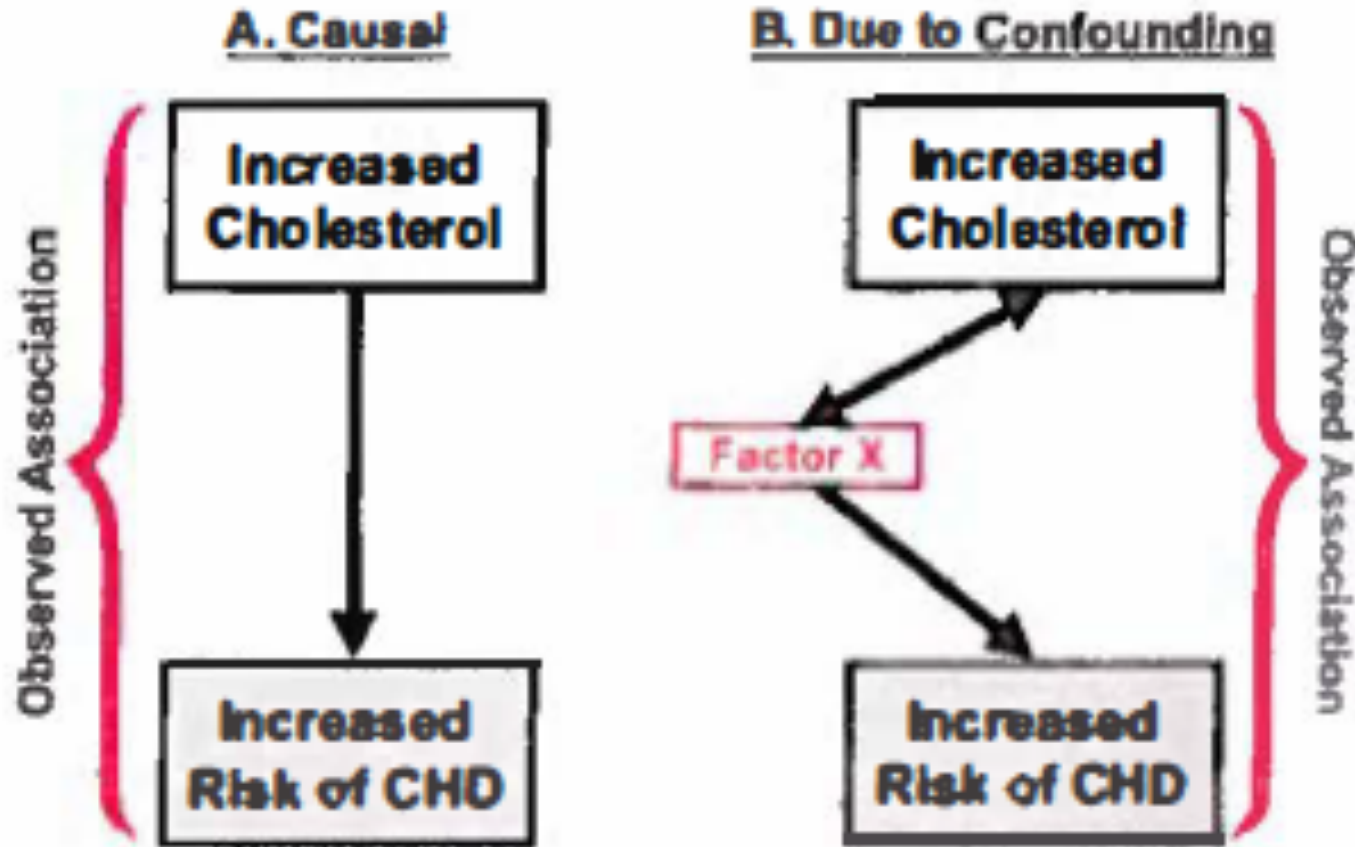


# Association between Coffee Drinking and Risk of Pancreatic Cancer



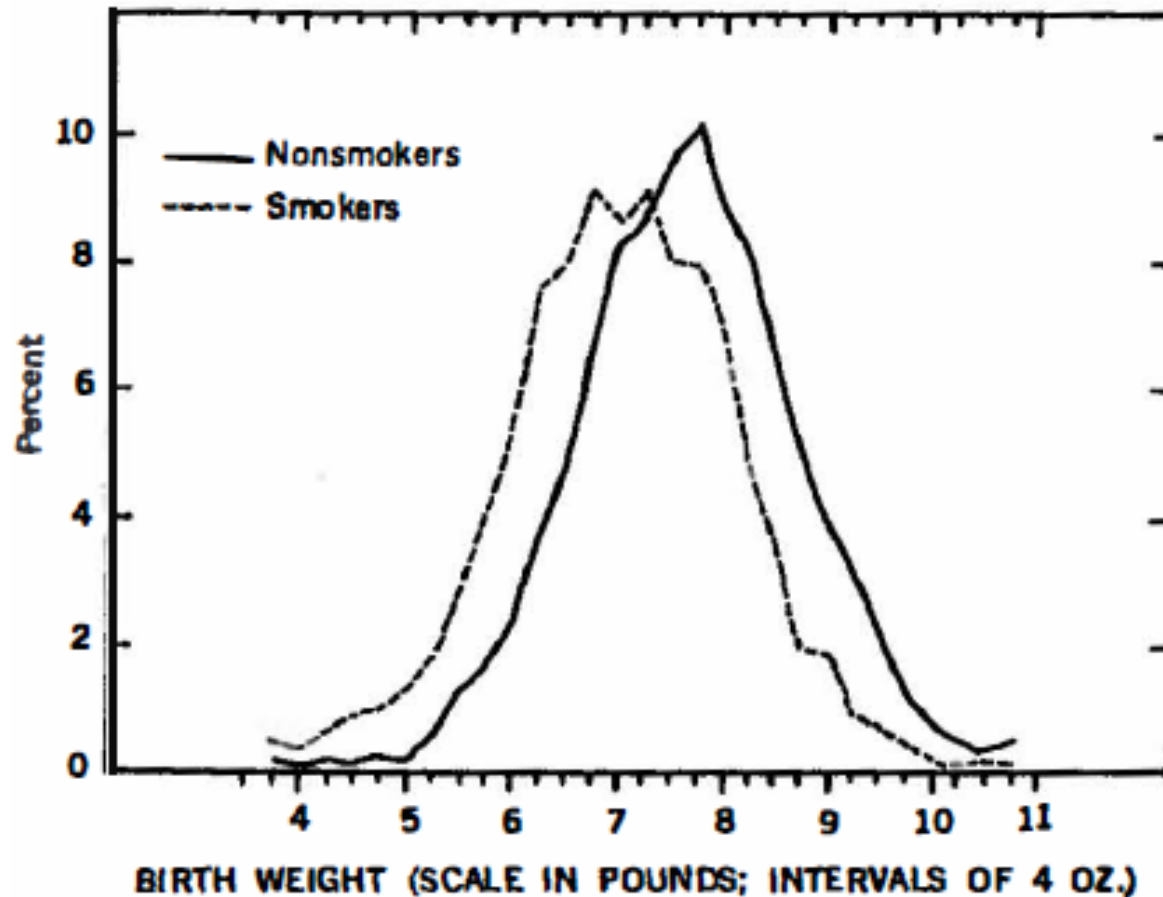


# Association between Cholesterol and Risk of Coronary Heart Disease (CHD)



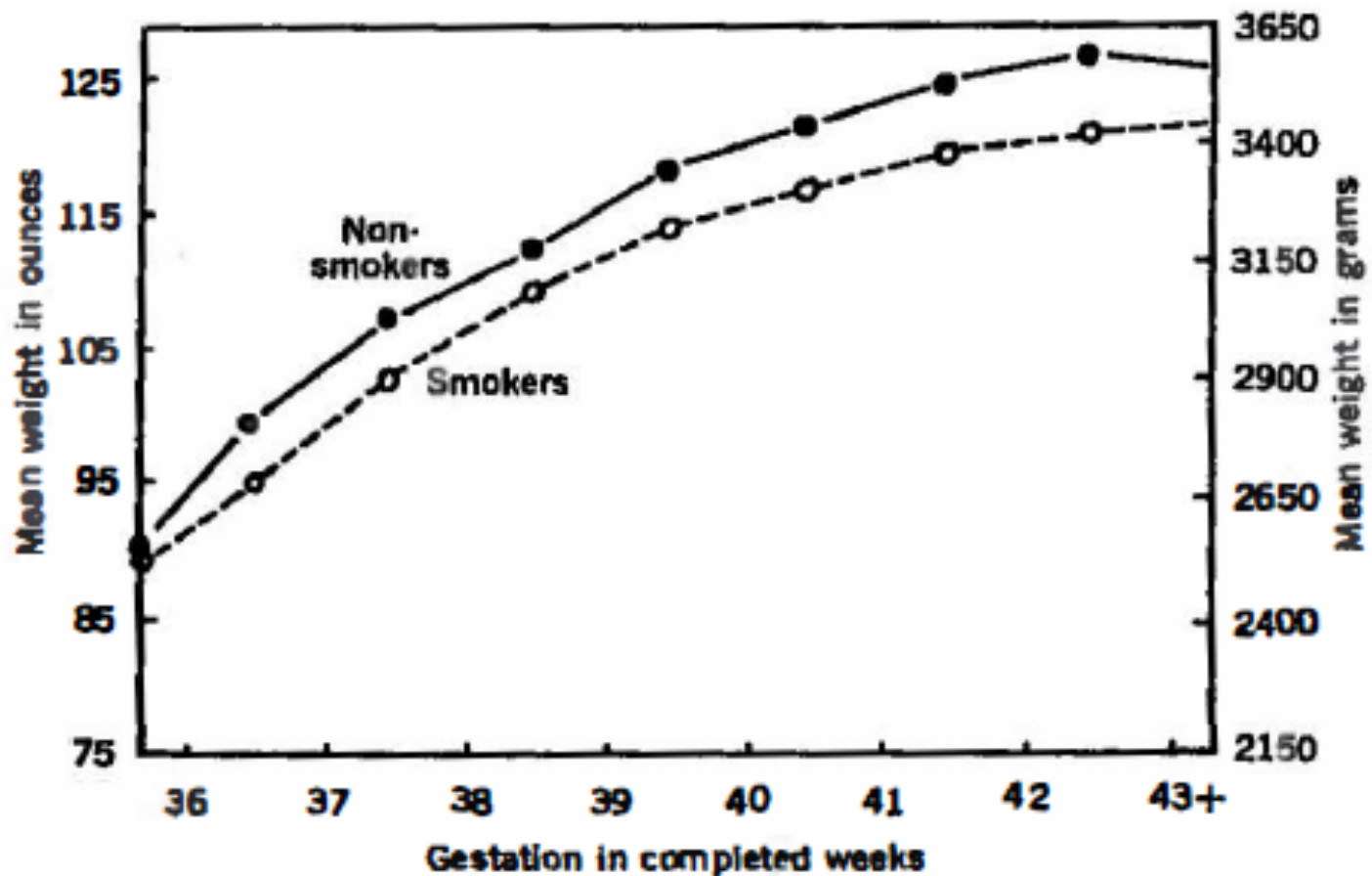


# Association between Cigarette Smoking by Pregnant Women and Low Birth Weight in Their Infants



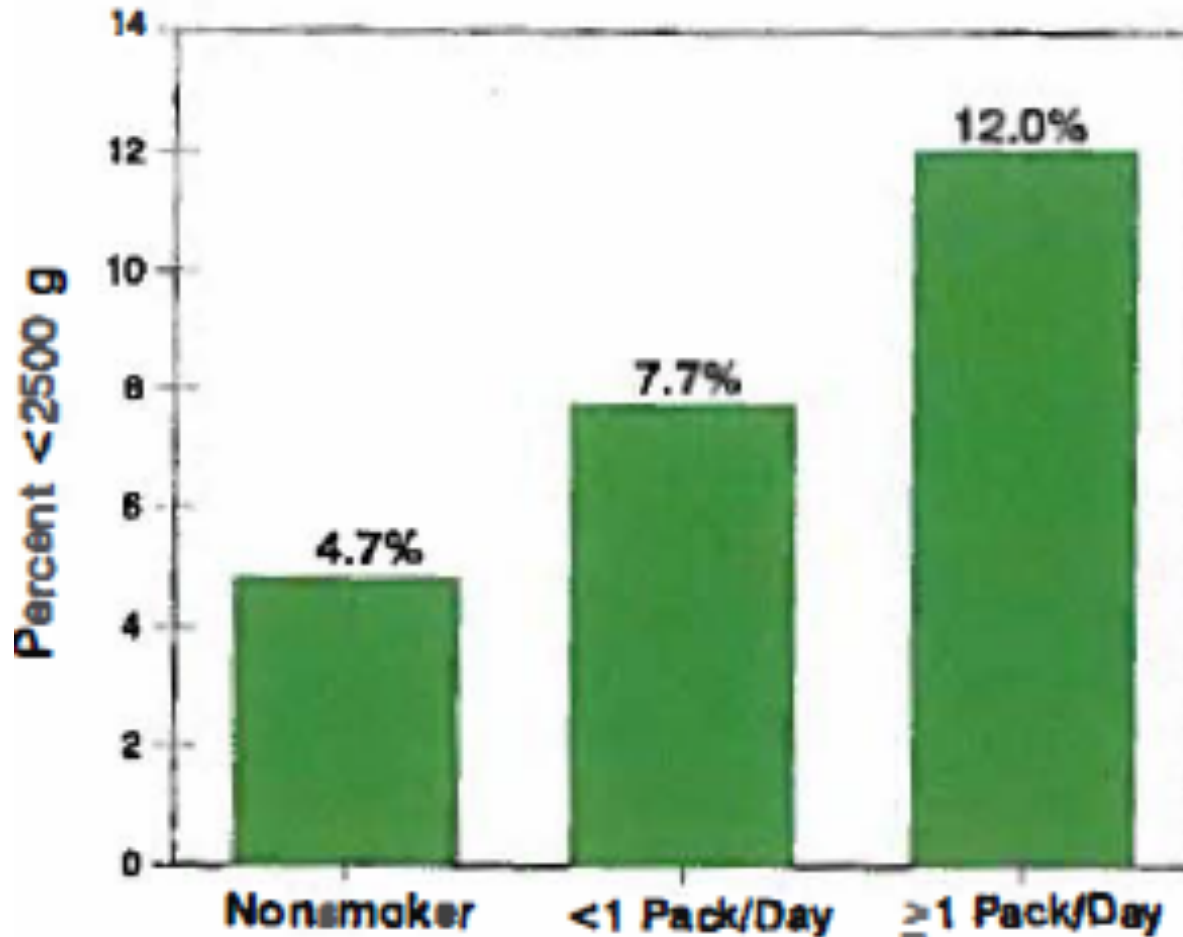


# Association between Cigarette Smoking by Pregnant Women and Low Birth Weight in Their Infants

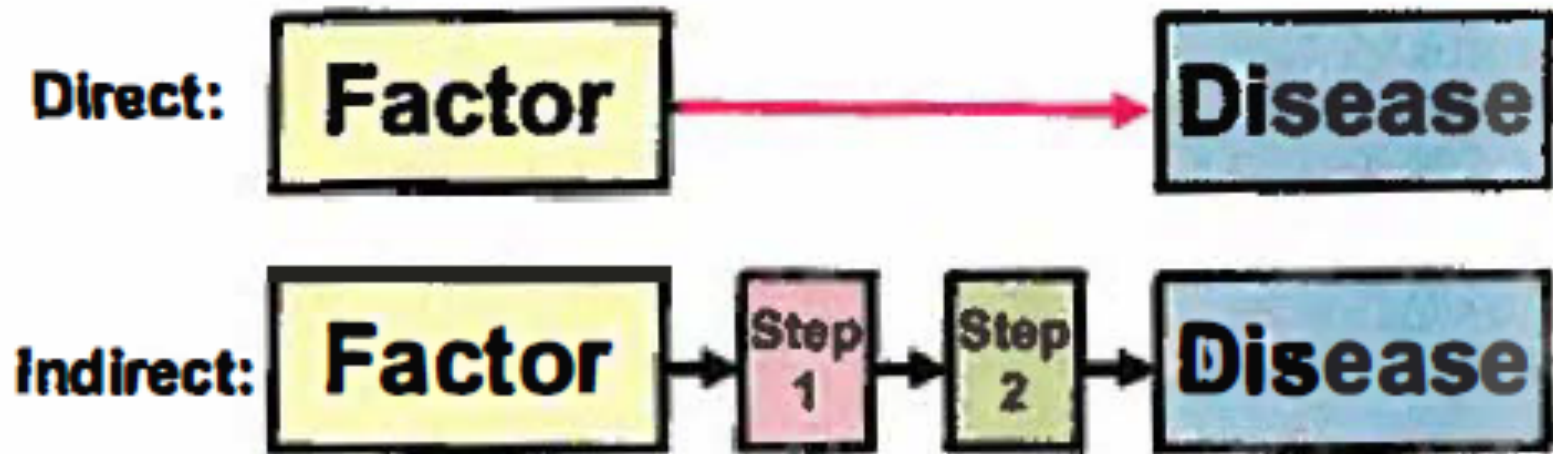




# Association between Cigarette Smoking by Pregnant Women and Low Birth Weight in Their Infants



# Types of Causal Relationships





# Types of Causal Relationships

- 1. Necessary and Sufficient**
- 2. Necessary but not sufficient**
- 3. Sufficient but not necessary**
- 4. Neither sufficient nor necessary**

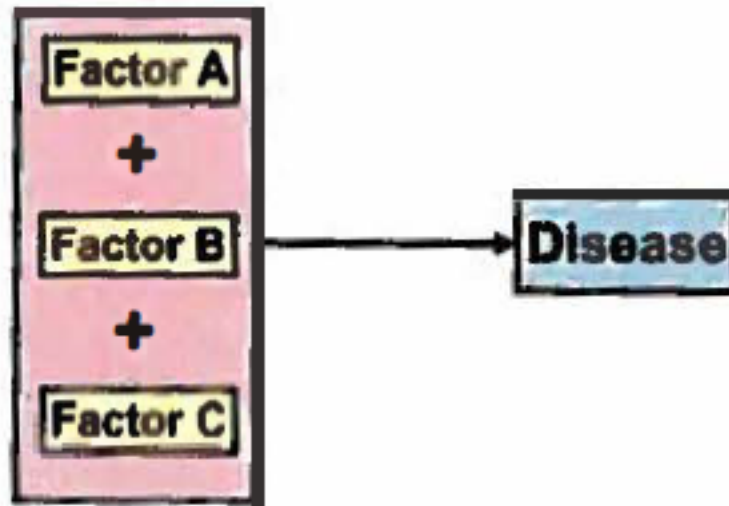
# 1. Necessary and Sufficient



**Without Factor A, the disease never develops (A is necessary), and in the presence of Factor A, the disease always develops (A is sufficient).**

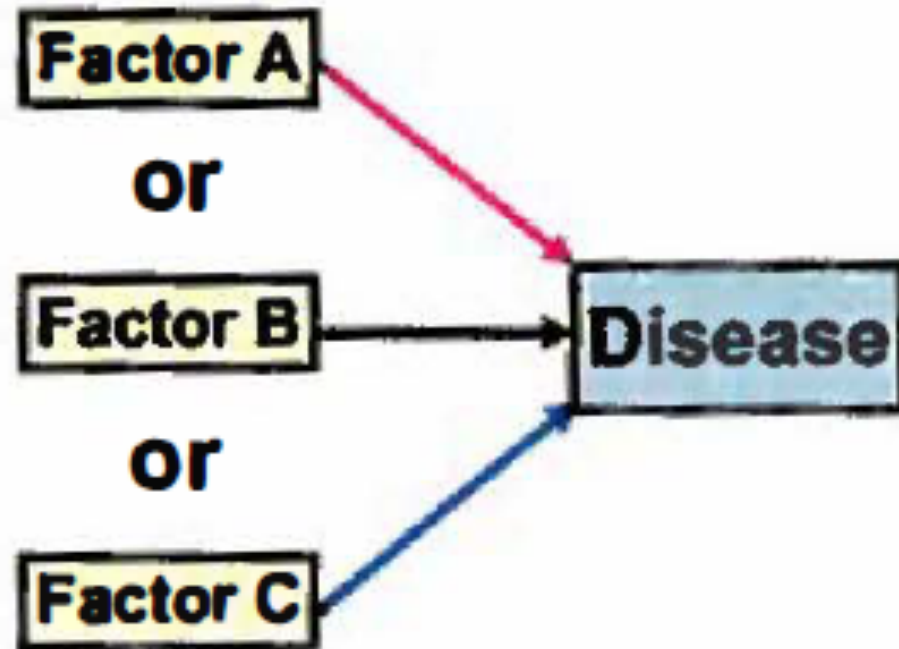
**e.g.: Electric shock---electrocute**

## 2. Necessary But Not Sufficient



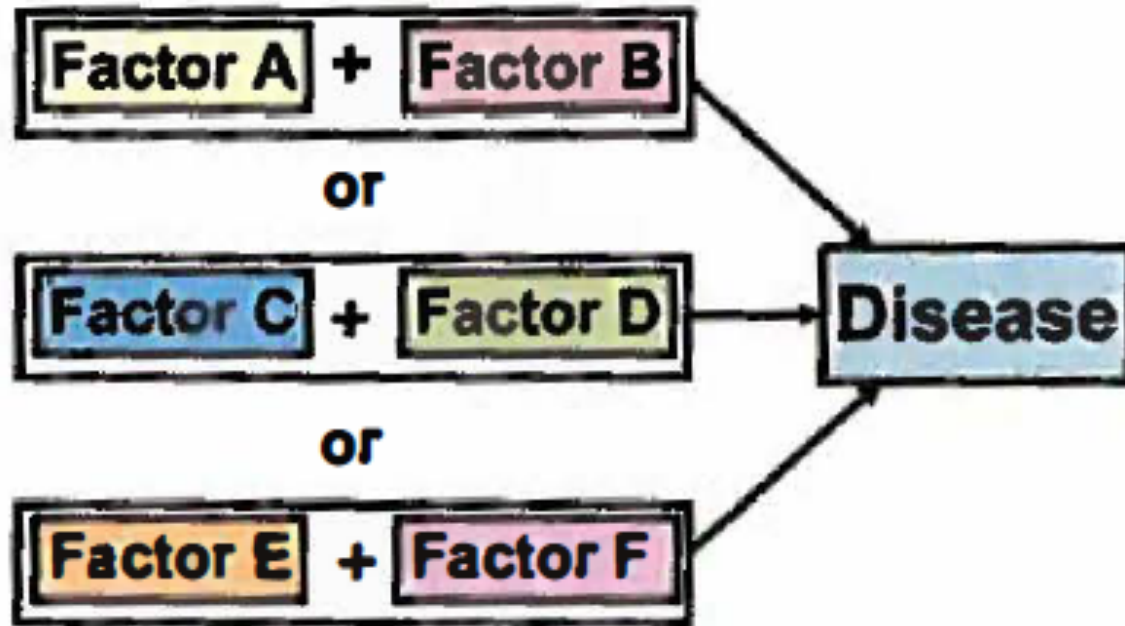
**Multiple factors are required (often in a specific temporal sequence) for the disease to develop.**

### 3. Sufficient But Not Necessary



**Factor A alone can produce the disease (A is sufficient), but so can other factors that are acting alone (A is not necessary).**

## 4. Neither Sufficient Nor Necessary



A factor, by itself, is neither sufficient nor necessary to produce the disease.

This model represents the causal relationships in most chronic diseases.



# Guidelines for Judging Whether an Observed Association is Causal

1. Temporal relationship
2. Strength of the association
3. Dose-response relationship
4. Replication of the findings
5. Biologic plausibility
6. Consideration of alternate explanations
7. Cessation of exposure
8. Consistency with other knowledge
9. Specificity of the association

**Example: Assessment of the evidence suggesting *Helicobacter pylori* (*H. pylori*) as a causative agent of duodenal ulcers**



# 1. Temporal relationship

The causal factor must have occurred before the disease developed.

*H. pylori* is strongly linked to chronic gastritis. About 11% of chronic gastritis patients go on to develop duodenal ulcers over a 10-year period.

# 2. Strength of the association

The stronger the association, the more likely it is that the relation is causal.

*H. pylori* is found in >90% of patients with duodenal ulcer.



### 3. Dose-response relationship

As the dose of exposure increases, the risk of disease also increases.

Density of *H. pylori* per square millimeter of gastric mucosa is higher in patients with duodenal ulcer than in patients without duodenal ulcer.

### 4. Replication of the findings

If the relationship is causal, we would find it consistently in different studies and in different populations.

Many of the observations regarding *H. pylori* have been replicated repeatedly.



## 5. **Biologic plausibility**

The hypothesized causal relationship should be coherent with the current body of biologic knowledge.

*H. pylori* has binding sites on animal cells and can follow these cells into the duodenum, where it induces inflammation and weakening of the mucosa.

## 6. **Consideration of alternate explanations**

It is important to consider whether the investigators have taken other possible explanations into account and the extent to which they have ruled out such explanations.

Data suggest that smoking can increase the risk of duodenal ulcer in *H. pylori*-infected patients but is not a risk factor in patients in whom *H. pylori* has been eradicated.



## 7. Cessation of exposure

If a factor is a cause of a disease, we would expect the risk of the disease to decrease when exposure to the factor is reduced or eliminated.

Eradication of *H. pylori* heals duodenal ulcers, and long-term ulcer recurrence rates are 0 after the eradication.

## 8. Consistency with other knowledge

If a relationship is causal, we would expect the findings to be consistent with other data.

Prevalence of *H. pylori* infection is the same in men as in women, and the incidence of duodenal ulcer is also equal in men and women.



## 9. Specificity of the association

An association is specific when a certain exposure is associated with only one disease. (This is the weakest of all the guidelines.)

*H. pylori* is not specific to duodenal ulcer. It is also found in patients with gastric ulcer and even in asymptomatic individuals.



# Modifications of the Guidelines for Causal Inferences

Stage 1: Categorization of the evidence by the quality of its sources (in descending order of quality.)

- 1) Trials
- 2) Cohort or case-control studies
- 3) Time-series studies
- 4) Case-series studies

Stage 2: Evaluation of the evidence of a causal relationship (in descending order of priority)

- 1) Major criteria: temporal relationship, biological plausibility, consistency, alternative explanations
- 2) Other considerations: dose-response relationship, strength of association, cessation effects



# 1. Trials

- a) Randomized, double-blind, placebo-controlled with sufficient power appropriately analyzed.
- b) Randomized, but blindness not achieved.
- c) Nonrandomized trials with good control of confounding, that are well conducted in other respects.
- d) Randomized, but with deficiencies in execution or analysis (e.g., insufficient power, major losses to follow-up, suspect randomization, analysis with exclusions).
- e) Nonrandomized trials with deficiencies in execution or analysis.



## 2. Cohort or case-control studies

- a) Hypothesis specified before analysis, good data, confounders accounted for.
- b) As above, but hypothesis not specified before analysis.
- c) Post hoc, with problems in the data or the analysis



### 3. Time-series studies

- a) Analyses that take confounding into account.
- b) Analyses that do not consider confounding

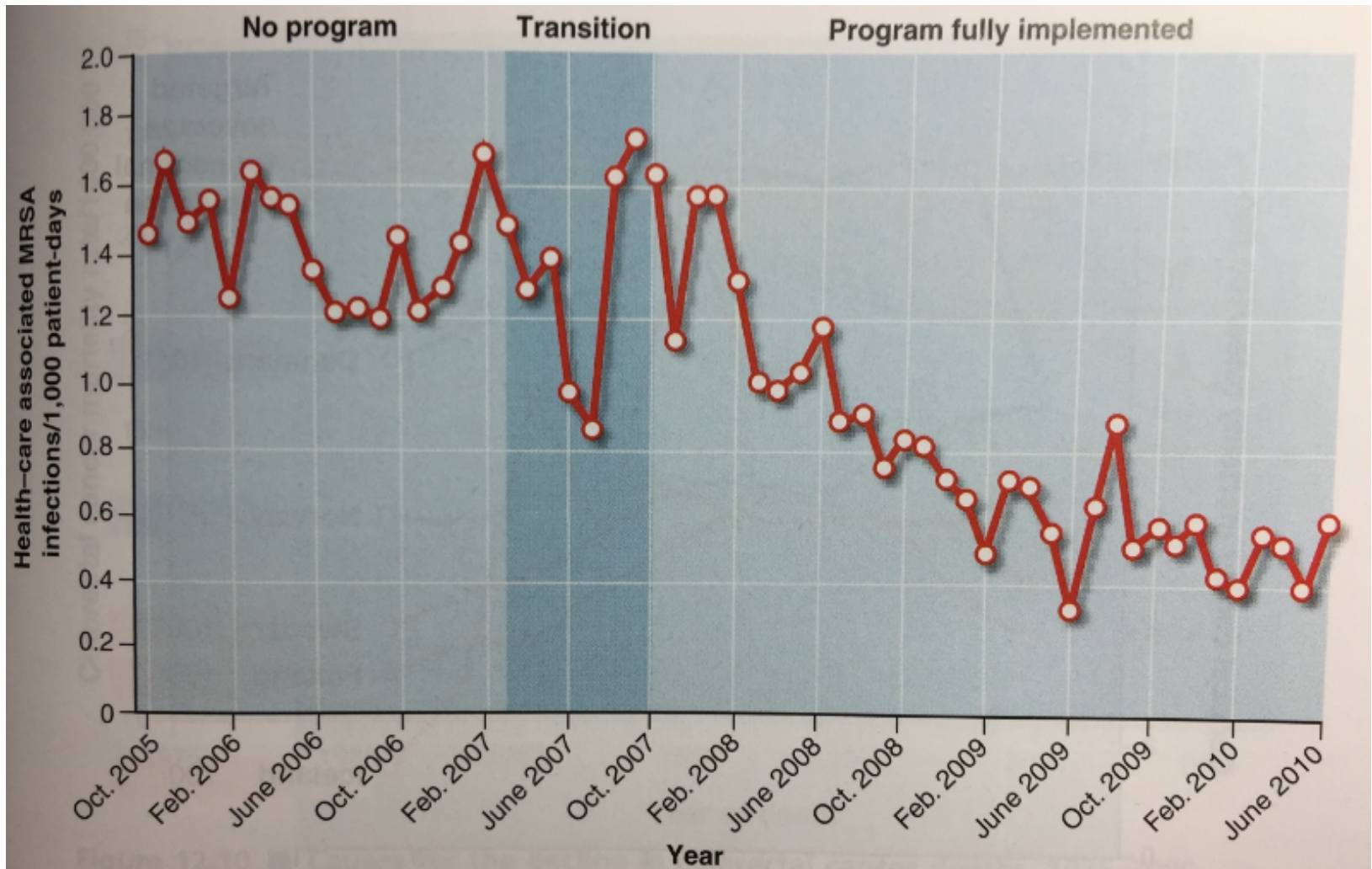
### 4. Case-series studies

Series of case reports without any specific comparison group.



# A time-series study.

Effect of a program to reduce methicillin-resistant Staphylococcus aureus(MRSA) infections in Veterans AFFAIRS FACILITIES





# Other Considerations in Reviewing Evidence

- Precision of definition of the outcome being measured
- Degree to which the study methodology has been described
- Adequacy of the sample size
- Degree to which characteristics of the study population and the intervention have been described
- Internal validity: well designed & well conducted
- External validity: generalizability



# Modifications of the Guidelines for Causal Inferences

Stage 1: Categorization of the evidence by the quality of its sources (in descending order of quality.)

- 1) Trials
- 2) Cohort or case-control studies
- 3) Time-series studies
- 4) Case-series studies

Stage 2: Evaluation of the evidence of a causal relationship (in descending order of priority)

- 1) Major criteria: temporal relationship, biological plausibility, consistency, alternative explanations
- 2) Other considerations: dose-response relationship, strength of association, cessation effects



# 1. Major criteria

- a) **Temporal relationship**: An intervention can be considered evidence of a reduction in risk of disease only if the intervention was applied **before** the time the disease would have developed.
- b) **Biological plausibility**: A biological plausible mechanism should be able to explain why such a relationship would be expected to occur.
- c) **Consistency**: Single studies are rarely definitive. Study findings that are replicated in different populations and by different investigators carry more weight than those that are not. If the findings are inconsistent, the inconsistency must be explained.
- d) **Alternative explanations (confounding)**: The extent to which alternative explanations have been explored is an important criterion in judging causality.



## 2. Other considerations

- a) **Dose-response relationship**: If a factor is indeed the cause of a disease, usually (but not always) the greater the exposure to the factor, the greater the risk of the disease. Such dose-response relationship may not always be seen because many important biologic relationships are dichotomous, and reach a threshold level for observed effects.
- b) **Strength of the association**: Usually measured by the extent to which the RR or OR depart from 1 (either above or below).
- c) **Cessation effects**: If a factor has an effect, then the effect should cease when it is removed from a population (unless it has some carry-over effect).



# Tobacco Smoking and Lung Cancer



**Smoking**  
**causes fatal lung cancer**



# Lung cancer

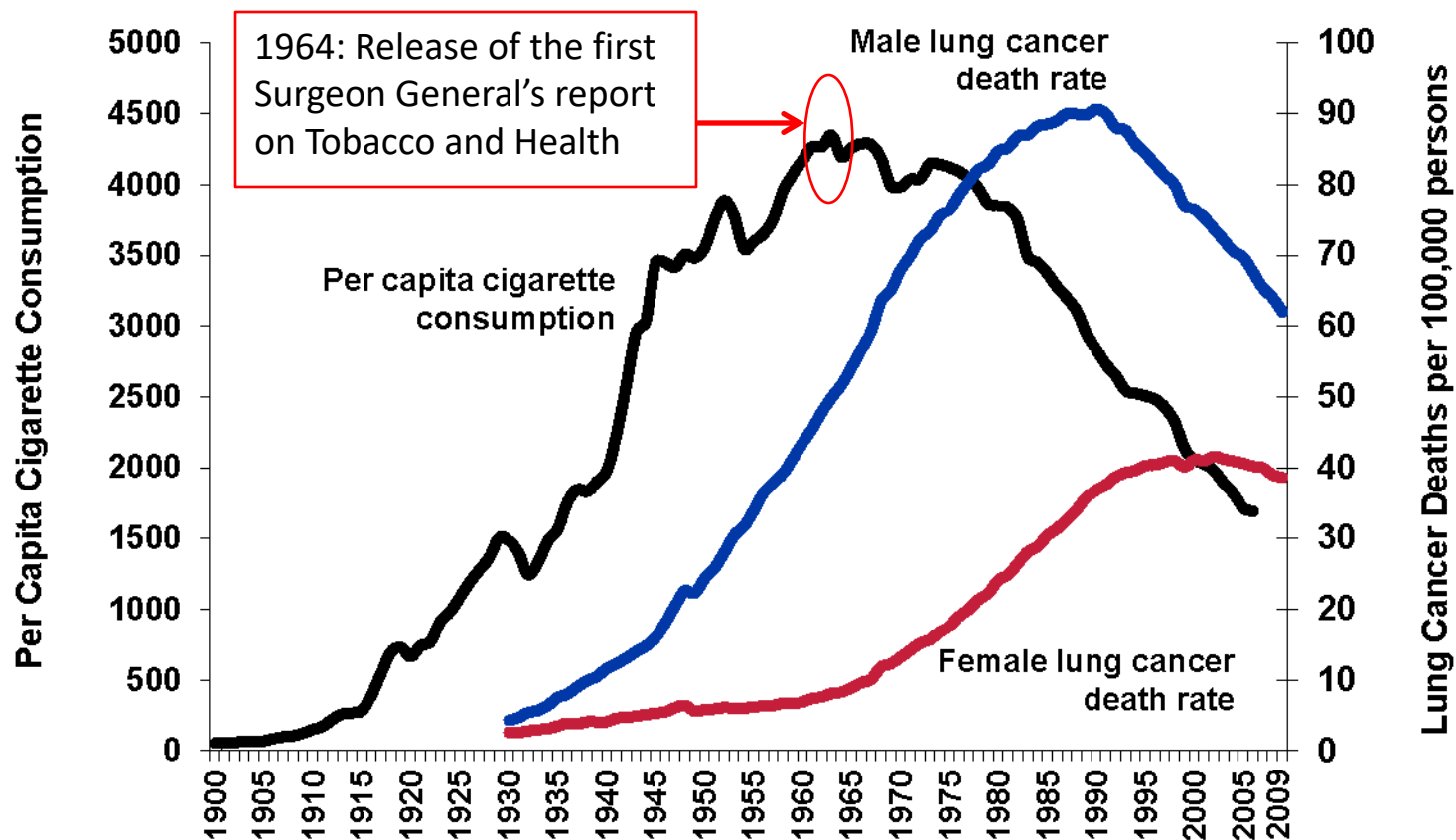
- Lung cancer has been the most common cancer worldwide since 1985, both in terms of incidence and mortality.
- In 2012, there were 1.83 million new lung cancer cases and 1.59 million deaths from lung cancer, which increased by 105% since 1985.
- 58% new cases occurred in less developed regions (as compared to 31% in 1980).



# Ecological Observation of Smoking Prevalence and Lung Cancer

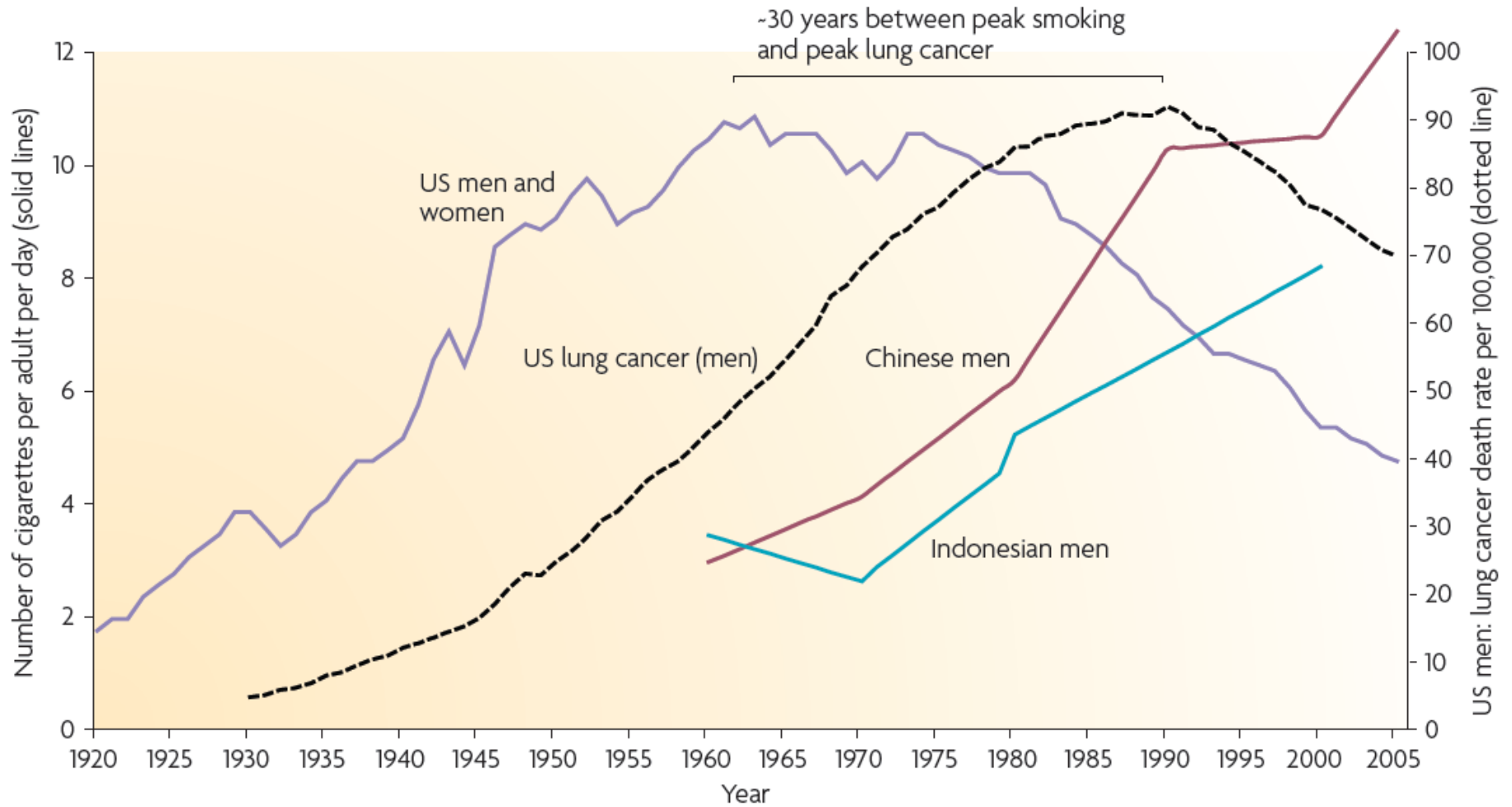
- The patterns of lung cancer occurrence determined largely by past exposure to tobacco smoking.
- Male lung cancer incidence started to decrease in 1980's and early 1990's, which is the direct result of the dramatic decrease in smoking prevalence during 1960s and 1970s.
- The rise in lung cancer incidence rates among women during 1970s and early 1980s and ended in the early 1990s is also the direct result of increase prevalence of smoking several decades ago.

# Trends in Tobacco Use and Lung Cancer Death Rates\* in the US



\*Age-adjusted to 2000 US standard population.

Source: Death rates: US Mortality Data, 1960-2009, US Mortality Volumes, 1930-1959, National Center for Health Statistics, Centers for Disease Control and Prevention. Cigarette consumption: US Department of Agriculture, 1900-2007.



**Figure 2 | Trends in cigarette consumption and male lung cancer rates, 1920–2005.** The number of cigarettes consumed per adult (males are used as the appropriate denominator for cigarette smoking as few females currently smoke in Asia) and the age-standardized lung cancer rates over time by country. The widespread automation of cigarette production in the early twentieth century turned cigarettes into a global commodity<sup>94</sup>. Lung cancer was a rare disease before the Second World War, and the large increase in lung cancer rates lagged behind consumption by three or more decades. The mean consumptions in Chinese men were 1, 4 and 10 cigarettes per day in 1952, 1972 and 1992, respectively, which were similar to the increases in cigarette consumption that were reported 40 years earlier in the United States (US)<sup>34,49</sup>. China has reported a marked increase in cigarette production since 2000. Increases in exposure to smoking at very young ages, combined with prolonged exposure, would be likely to increase the age-specific death rates in the future in China, Indonesia and other countries.



# Lung Cancer: Time Trends

- Trends in lung cancer incidence and mortality reflect the maturity of the smoking epidemic in different countries.
- Study of time trends in lung cancer incidence or mortality by age group shows that the level of risk is closely related to the birth cohort, or the cohort specific incidence is related to the smoking habit of the same generation.



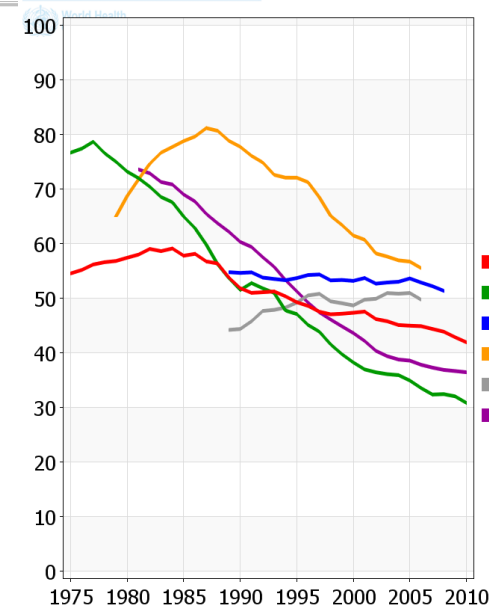
## Lung Cancer: Time Trends

- In countries where smoking was first established, decreased smoking prevalence leads to a decreased lung cancer rate such as in U.K., Finland, Australia, the Netherlands, New Zealand, the US, Singapore, and more recently Denmark, Germany, Italy, and Sweden.
- In Central and Eastern European countries (as Bulgaria, Hungary, Poland, Romania, the Czech Republic) – after increasing trends observed up to the late 1990s – lung cancer showed appreciable declines over the most recent decades, with particularly favorable patterns in young men.
- Lung cancer rates in women are generally increasing since tobacco habit is fairly recent, except in countries where smoking prevalence among women is declining.

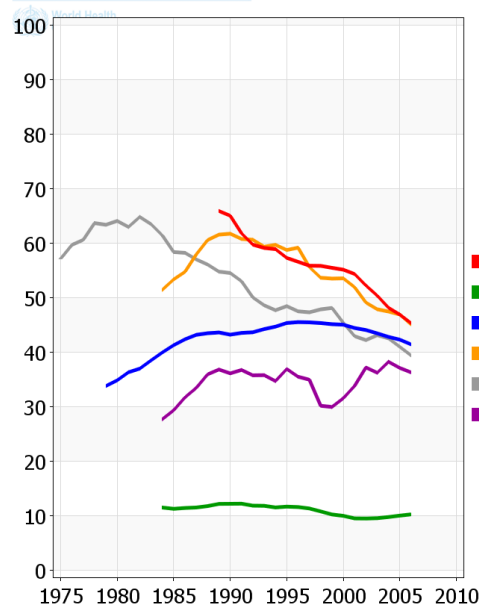
# Trends of Incidence of Lung cancer



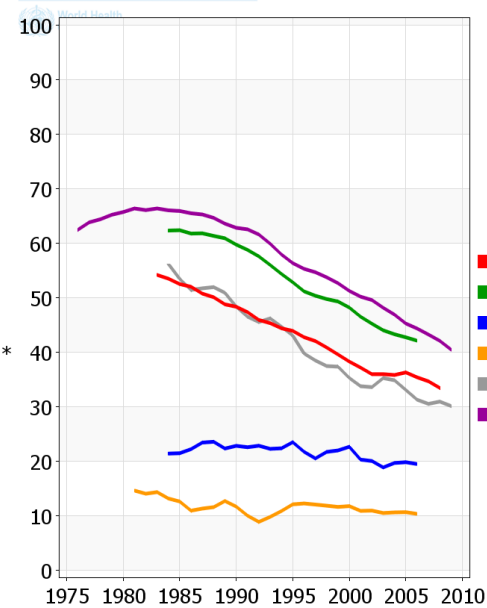
**Men**



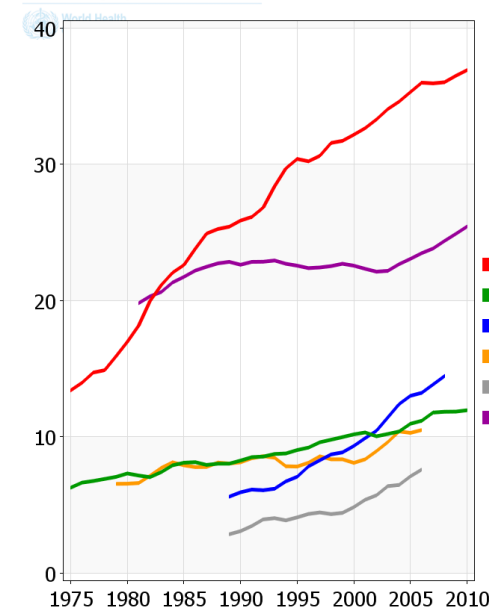
International Agency for Research on Cancer



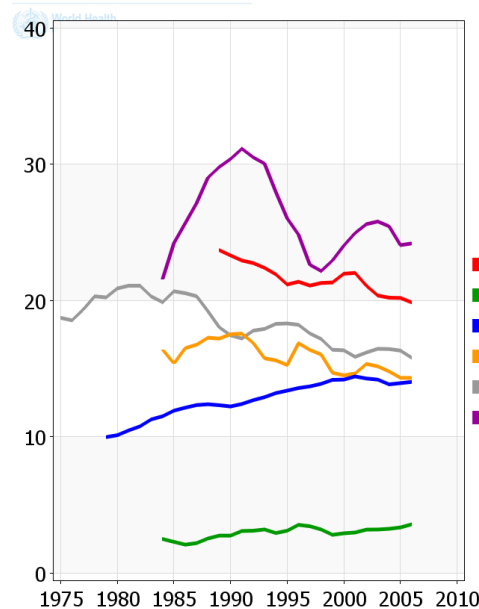
International Agency for Research on Cancer



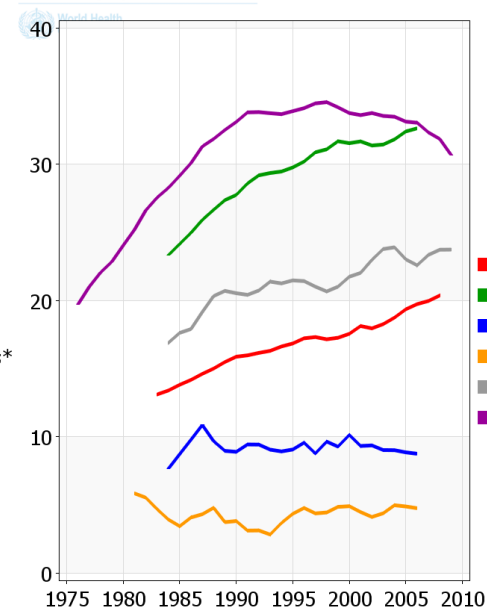
**Women**



International Agency for Research on Cancer



International Agency for Research on Cancer



\*Regional Data



# Lung Cancer: Cigarette Smoking

- Cigarette smoking is more strongly associated with lung cancer than with any other cancer site.
- Cigarette smoking caused about 71% of lung cancer deaths worldwide and 1 in 9 smokers eventually develops lung cancer.
- It accounts for 87% of lung cancer deaths in men and 70% in women in the US.
- The risk of developing lung cancer is affected by the number and type of cigarettes smoked and the numbers of years a person has been smoking.
- The risk of lung cancer is as high in smokers of “light” or “low-tar” yield cigarettes as in those who smoke “regular” or “full-flavored” products.



# Lung Cancer Histological Types and Smoking

- Lung cancer arises from the cells of the respiratory epithelium
  - Non-small cell lung cancer: adenocarcinoma (38.5%), squamous cell carcinoma (20%), and large cell carcinoma (2.9%)
  - Small cell lung cancer (15%)
- Adenocarcinoma is more common in women than men in both smokers and non-smokers.
- The incidence of adenocarcinoma has increased greatly in the past few decades, which may be attributable to the deeper inhalation of lowered tar-containing and nicotine containing cigarettes, leading to a more peripheral distribution of cigarette smoke in the lung.
- Smoking is related to all histological types of lung cancer.
- The relative risk is greater for squamous-cell and small-cell carcinomas than for adenocarcinomas.



# First Case-Control Studies

- In 1950, five case-control studies published within a 9 month period reported an association between cigarette smoking and lung cancer in men.
- Two largest and most influential studies are
  - A British case-control study included 1732 lung cancer cases and 743 controls (Doll & Hill, *BMJ* 1950)
  - An American case-control study included 605 lung cancer cases and 780 controls (Wynder & Graham, *JAMA* 1950)
- Subsequently, the Surgeon General of the United States concluded in the 1964 report that “cigarette smoking is the major cause of lung cancer.”



# The Consistency of the Association

- In 1984 surgeon general's report, 29 retrospective and 7 prospective studies have all demonstrated an association between cigarette smoking and lung cancer.
- Similar findings had thus reported from a large number of studies by different investigators, using different methodological approaches, and in different populations.



# Biological Credibility

- Over 60 carcinogens are found from cigarette smoke.
- Carcinogens may cause DNA damage.
- Mutation of Tumor suppresser genes are linked to cigarette smoking.
- Animal studies also support the relationship between cigarette smoking and lung cancer.



# The Strength of the Association

Table 1. Relative risk of lung cancer and major subtypes by smoking status

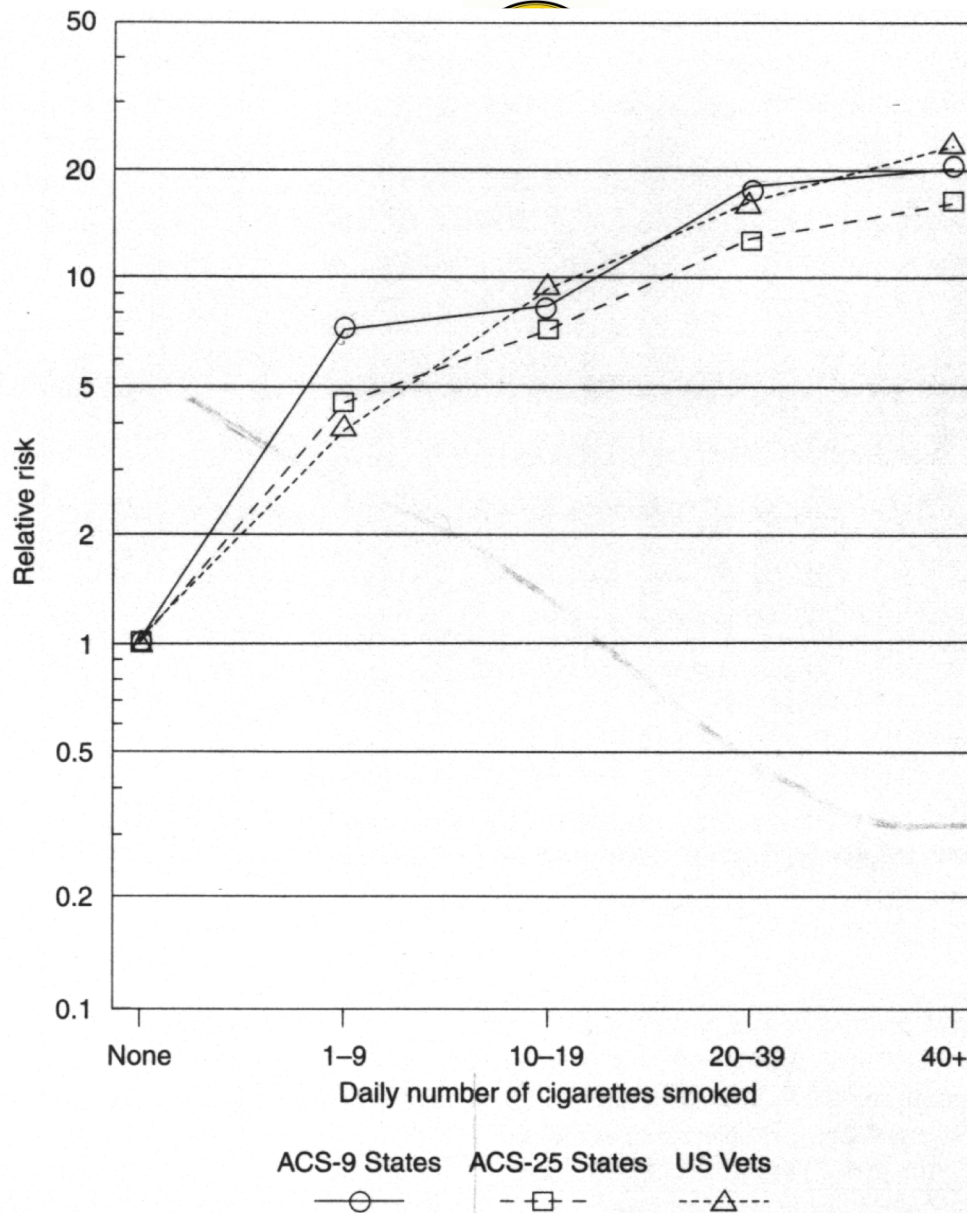
Studies	Gender	Smoking cigarettes	Controls	All histologies		Squamous cell cancer		Small cell lung cancer		Adenocarcinoma	
				Cases	OR <sup>1</sup> (95% CI)	Cases	OR <sup>1</sup> (95% CI)	Cases	OR <sup>1</sup> (95% CI)	Cases	OR <sup>1</sup> (95% CI)
All	Men	Never	2,83	220	1.0	51	1.0	22	1.0	99	1.0
		Former	5,647	3,496	7.5 (6.5–8.7)	1,585	14.7 (11.0–9.6)	460	10.1 (6.5–15.5)	862	4.2 (3.4–5.2)
		Current	3,829	6,784	23.6 (20.4–27.2)	3,038	45.6 (34.3–60.6)	1,249	45.7 (29.9–70.0)	1,398	10.8 (8.7–13.3)
		Other tobacco	399	153	5.9 (4.6–7.4)	73	12.6 (8.6–18.4)	29	9.6 (5.4–17.0)	25	2.0 (1.3–3.2)
		Total	12,758	10,651		4,747		1,760		2,384	
	Women	Never	1,902	609	1.0	96	1.0	50	1.0	343	1.0
		Former	657	485	2.8 (2.4–3.3)	124	4.9 (3.7–6.6)	60	4.2 (2.8–6.2)	208	2.0 (1.6–2.4)
		Current	691	1,423	7.8 (6.8–9.0)	343	13.6 (10.5–17.7)	331	21.7 (15.5–30.1)	462	4.2 (3.5–5.0)
		Total	3,250	2,517		563		441		1,013	

Pooled analysis included 13,169 cases and 16,010 controls from Europe and Canada (Pesch B, IJC 2012).

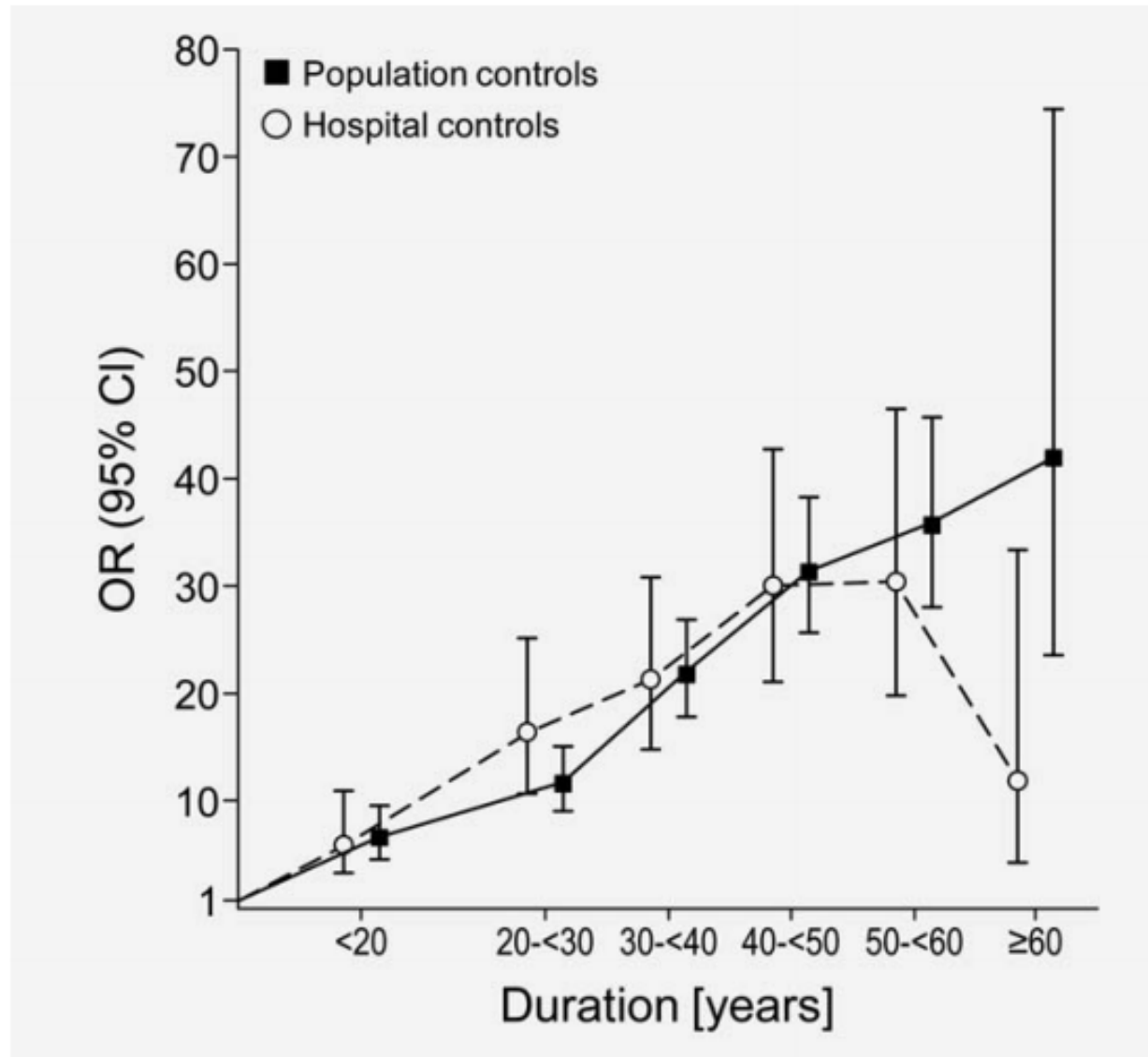


# Dose-Response Relationship

- The existence of a dose-response relationship between smoking and lung cancer constitutes further evidence of a causal relationship.
- Many studies have demonstrated a dose-response relationship between cigarette smoked per day, years of smoking and pack-years of smoking and lung cancer.



**Figure 9-9** Data from three cohort studies showing the relation between lung cancer risk and daily number of cigarettes smoked. *Source:* Adapted with permission from *Tobacco Smoking*, IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, No. 38, pp. 215-216, © 1986, International Agency for Research on Cancer.



**Figure 1.** Duration-response trend of the adjusted lung cancer relative risks in studies with population controls vs. studies with hospital controls in male current smokers.

(Pesch B, IJC 2012)



## Specificity of the Association

- The association between smoking and lung cancer has “a high degree of specificity”.
- Although smoking does not always result in lung cancer and lung cancer has other causes, it is estimated that 85% to 90% of lung cancer deaths in men is due to cigarette smoking.
- The observation that all smokers do not develop lung cancer indicate that there is a genetic component to the pathogenesis of lung cancer or other competing causes of death.



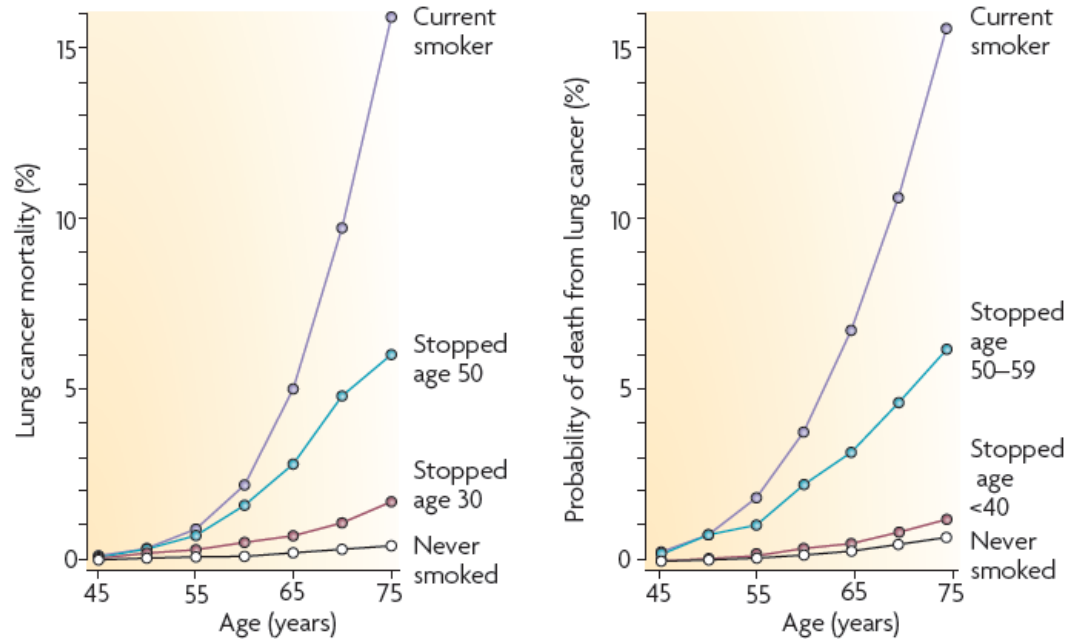
# Temporal Relationship

- A temporal relationship between smoking and lung cancer is supported by several prospective cohort studies when exposure to cigarette smoking was obtained before the occurrence of lung cancer.
- The latent period between first exposure and the onset of the disease is about 20-30 years.



# Coherence of the Association

- Distribution of the disease and of the cigarette smoking are similar with regards to person, place, and time.
- Rapid rise in cigarette consumption resulted in the rapid rise of lung cancer after a period of latency.
- The distribution of cigarette smoking among sexes, among rural and urban residents, and among different socioeconomic classes was similar to the distribution of lung cancer among these groups.
- Those who quit smoking had a decreased risk.



**Figure 3 | Risk of death from lung cancer in smokers and ex-smokers.** The age-specific probability of death from lung cancer for the United Kingdom and United States, stratified by continued smoking or cessation at various ages. The United Kingdom data (a) are from retrospective studies<sup>18</sup>, courtesy of R. Peto and J. Boreham. For each age, the relative risks match those in a case-control study of smoking, and an appropriately weighted average of the absolute risks matches the national lung cancer death rates. The United States data (b) are from the American Cancer Society prospective study of 1.2 million subjects during the first 10 years of follow up, courtesy of M. Thun<sup>29</sup>. These data omit the earlier years of follow up (1981–1983). The same study showed that few of those who stop would restart smoking. In both of these studies, those who stopped smoking did so within 5 years of the stated age, and almost all smokers had used cigarettes. Similar results on reductions in lung cancer risk in ex-smokers are seen in Poland<sup>18,95</sup>. The excess lung cancer mortality avoided in men who stopped smoking by age 40 was 91% in Germany and 80% in Italy<sup>96</sup>. A minority of lung cancers are not due to smoking, and the United States rates of lung cancer that are not due to smoking have changed little from the 1960s to 1990s<sup>6,46,97</sup>. A range of genetic factors has recently been suggested that modestly predicts lung cancer risk in smokers and non-smokers<sup>98</sup>. However, it is improbable that marked shifts in genetic susceptibility have occurred and, even if they had, such shifts would be unlikely to explain the dramatic changes in lung cancer that have been seen over a few years or decades<sup>26,29</sup>.



# Lung Cancer: Other Tobacco

- Cigar and pipe smoking have both been associated with increased risk for lung cancer.
- The early observation that cigar- or pipe-smokers had a lower risk for lung cancer than cigarette smokers is due to the lesser amount of tobacco smoked and the lower degree of inhalation.
- Smokers of other types of tobacco (e.g. bidis in India) are at similar risk to smokers of cigarette.



# Lung Cancer: Other Risk Factors

- Passive smoking may increase 30-50% of lung cancer risk.
- Family history of lung cancer may also related to lung cancer risk, which is probably because of polymorphisms of genetic metabolic genes.
- Chinese female non-smokers have a high risk for lung adenocarcinoma, which may be related to exposure to cooking fumes, indoor smoky coal emission, or air pollution.



# Lung Cancer: Other Risk Factors

- Occupational exposures to asbestos, some metals (nickel, arsenic and cadmium), radon (among miners) and ionizing radiation.
- Diet high in vegetables and fruits can protect against lung cancer. However, although intakes of carotenoids is associated with decreased risk, beta-carotene does not prevent lung cancer in 3 intervention trials.
  - The Alpha-Tocopherol, Beta Carotene Cancer Prevention (ATBC) Study: 29,133 male smokers aged 50 to 60 years in Finland.
  - The Beta-Carotene and Retinol Efficacy Trial (CARET): 18,314 men and women at increased risk for lung cancer.
  - The Physicians' Health Study: 22,071 male physicians; 11% were current smokers and 39% former smokers.

# Prospects for Prevention

- **Encourage smoking cessation among smokers.**
  - Disseminating health information
  - Mandating smoke-free areas
  - Implementing tax increases
- **Reinforcing social pressure to make smoking unattractive and difficult for adolescents to start smoking.**
- **Early detection of early lesions, by x-ray or other means.**



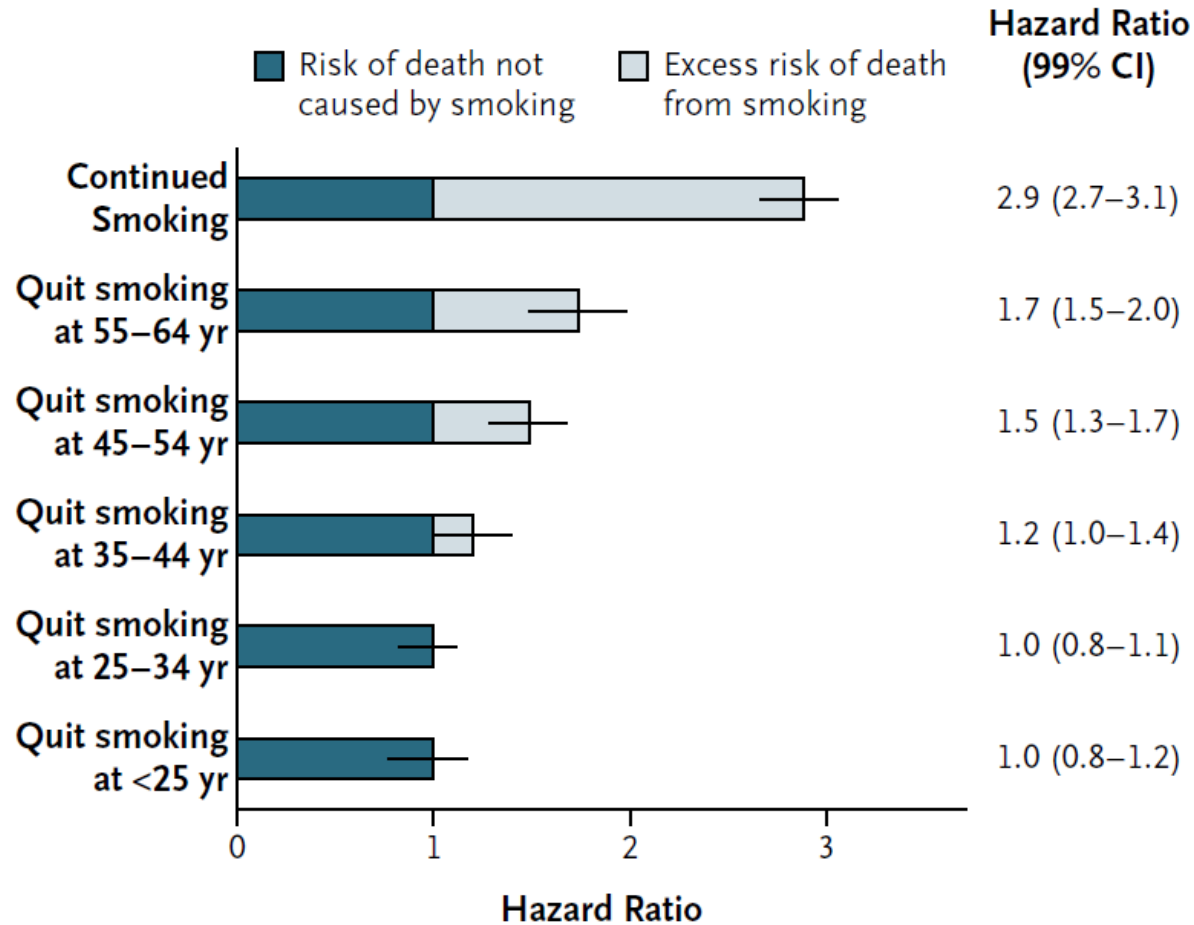


# Paradigm for Tobacco Control

- ▼ Minor's access
- ▼ Advertising

- ▲ Cessation activities
- ▲ Prevention activities
- ▲ Clean indoor air
- ▲ Regulation/Liability
- ▲ Price/Economic

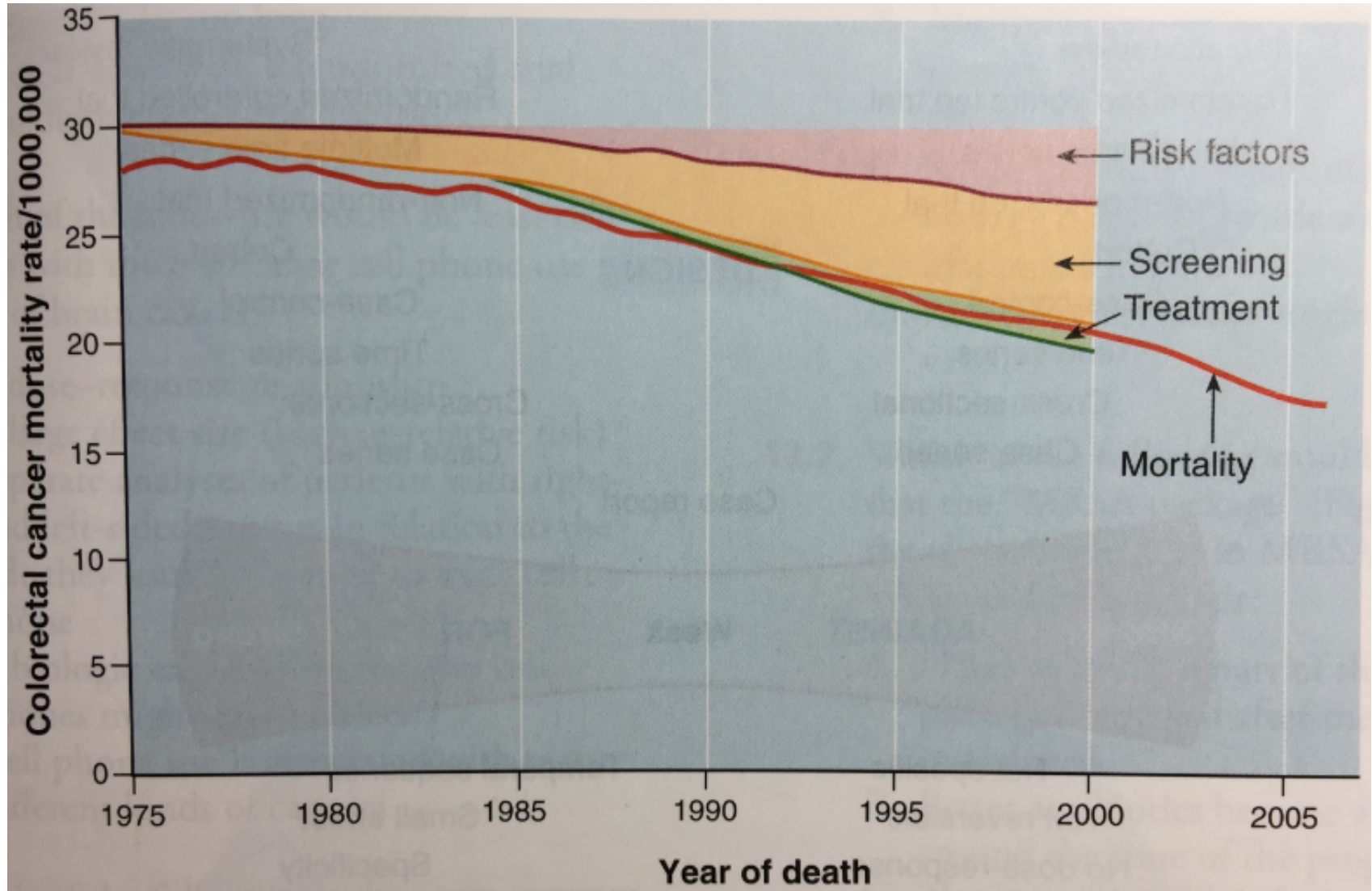




**Figure 4.** Risks of Death for Participants Who Continued to Smoke and for Those Who Quit Smoking According to Age at the Time of Cessation. The total and excess risks of death are shown for NHIS participants who continued smoking, as compared with those who quit smoking. CI denotes confidence interval.

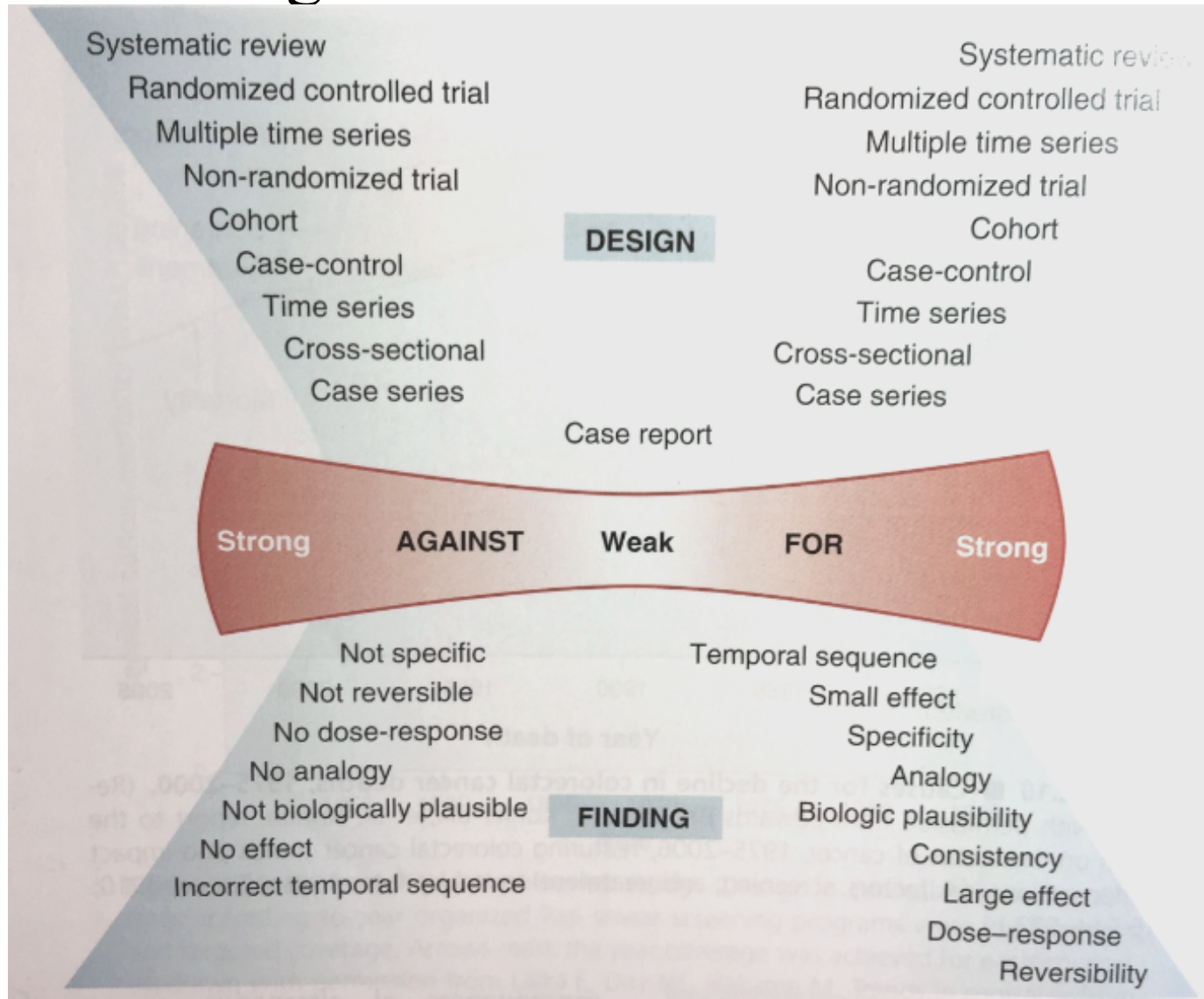


# Causes of the decline in colorectal cancer deaths, 1975-2000





# Relative strength of evidence of evidence for and against a causal effect





# Conclusion

When making decisions about causation, the preceding guidelines must be coupled with reasoned judgment about the entire body of available evidence.

我图网 oopic.com  
NO.20090602221387

**Thanks for your attention!**