Public Health Sciences 310
Epidemiologic Methods

Lecture 2
Study Design

January 9, 2024

Brian C.-H. Chiu

Today's Outline

- Epidemiologic Study Designs
- Surveillance
  - Age, period, and cohort effects
- Case-Control Studies
- Cohort Studies
  - Nested Case-control Studies
  - Case-cohort studies

Epidemiologic Study Designs

(Two types)

Qualitative

Descriptive: all are observational

"generate" hypothesis

hypothesis generating

Distribution

Clear, specific and measurable: Time, Place, Person

> Surveillance
> Ecologic
> Cross-sectional

> Case Report
> Mortality Study

Quantitative

Analytic: test a hypothesis

"Analyze" hypothesis

Determinants

Comparison group to measure causal effect

> Observational
  - Cohort
  - Case-Control
> Experimental
  - Clinical Trials/Interventions
  - Community Interventions

Today's Outline

- Epidemiologic Study Designs
- Surveillance
  - Age, period, and cohort effects
- Case-Control Studies
- Cohort Studies
  - Nested Case-control Studies
  - Case-cohort studies

Surveillance Studies: Rates & trends

The ongoing systemic collection, analysis and interpretation of health data essential to the planning, implementation and evaluation of health practice.

- Data Sources
  - Death certificates, birth certificates
  - Census data
  - Administrative data
  - Registries (e.g., SEER)
  - … Goal: To determine if there is a trend in incidence that is caused by an epidemic.

Descriptive Study Example

Leukemia & Lymphoma, 2013; Early Online: 1–7

ISSN: 1042-8194 print / 1029-2403 online
DOI: 10.3109/10428194.2012.760041

informa
healthcare

ORIGINAL ARTICLE: CLINICAL

An upward trend in the age-specific incidence patterns for mantle cell lymphoma in the USA

Briseis Aschebrook-Kilfoy[1], Donne Bennett D. Caces[2], Nicholas J. Ollberding[1], Sonali M. Smith[2] & Brian C.-H. Chiu[1,3]

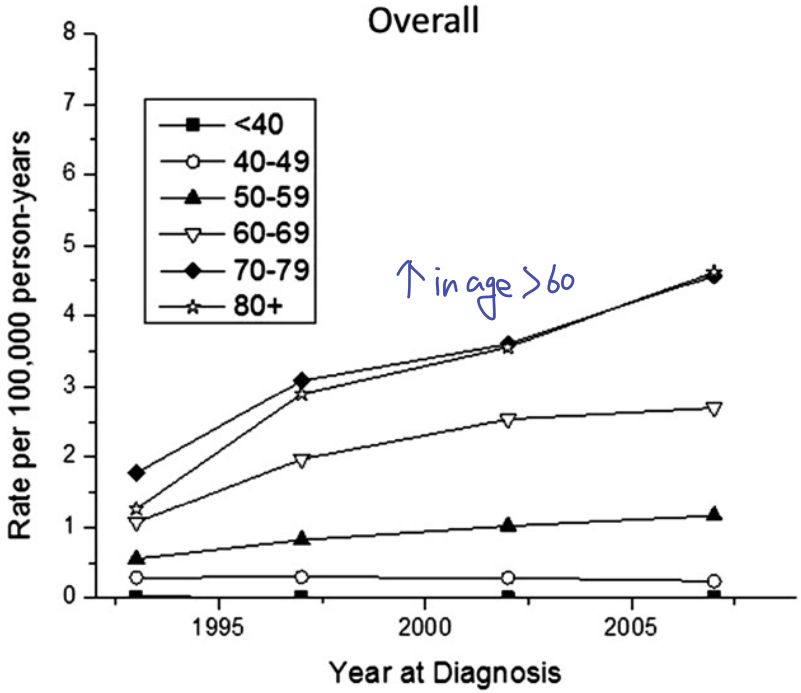
[1]*Department of Health Studies and* [2]*Division of Hematology/Oncology, Department of Medicine, University of Chicago, Chicago, IL, USA and* [3]*University of Chicago Comprehensive Cancer Center, Chicago, IL, USA*

Abstract

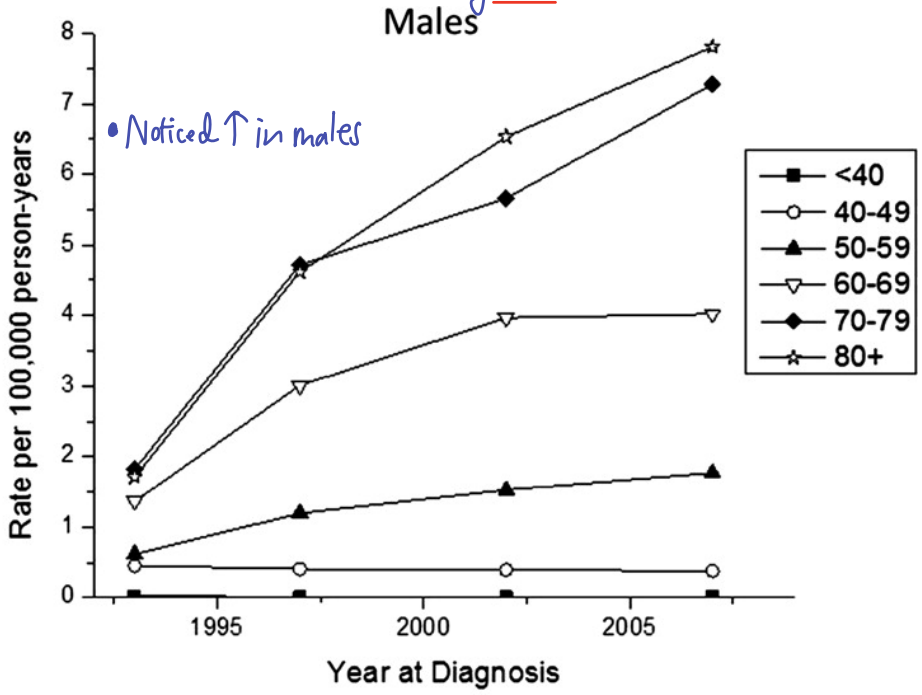
Although an increased incidence of mantle cell lymphoma (MCL) has been reported, age-specific incidence patterns have not been described. Further analyses could inform investigation into the etiology of this disease. We conducted an epidemiologic study using the 13 Surveillance, Epidemiology, and End Results (SEER) registries to evaluate MCL incidence from 1992 through 2009. We calculated the proportional changes in the incidence of MCL for subpopulations defined by age, race/ethnicity and gender over time and the racial/ethnic and gender disparities.

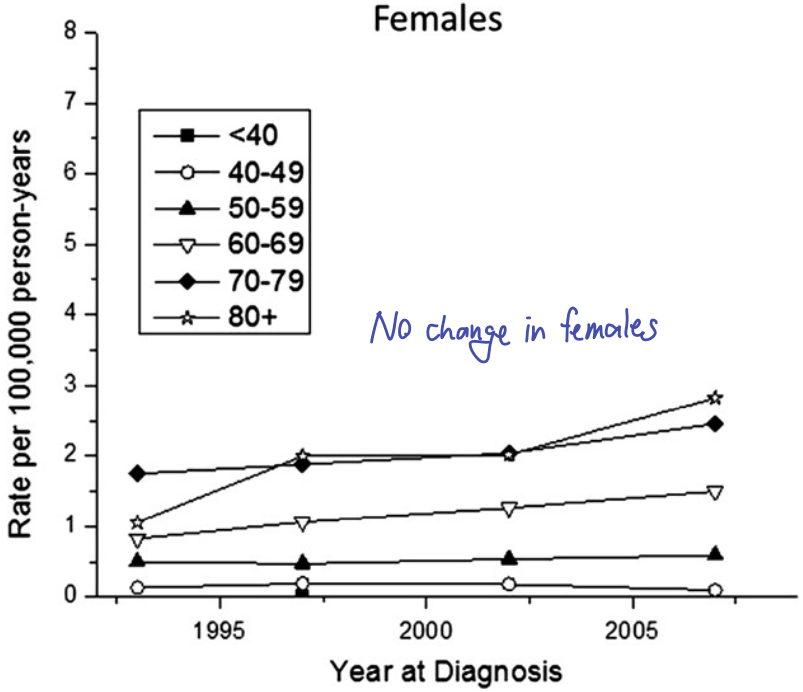
Sudden increase in MCL

change in diagnosis criteria?

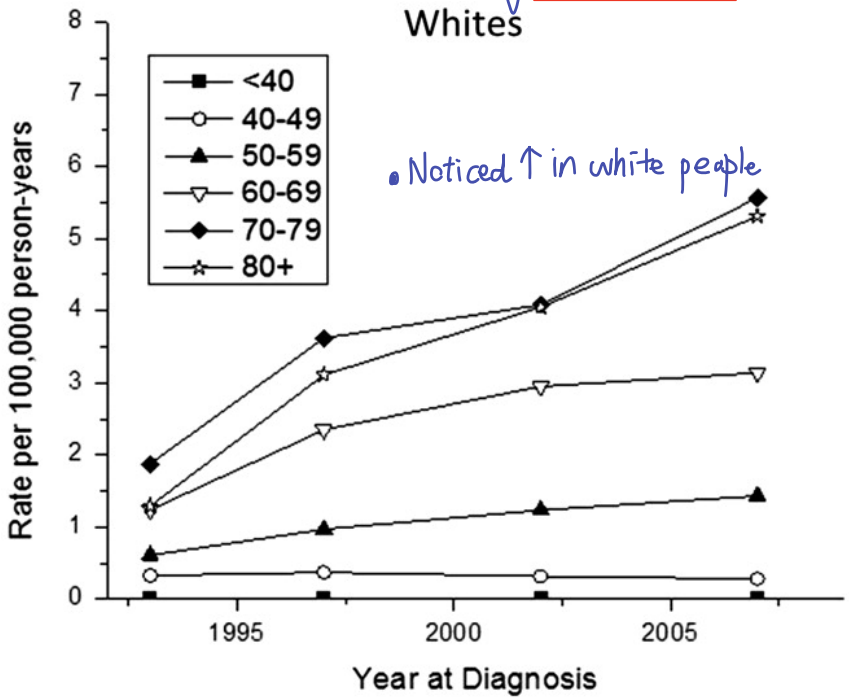


Stratified by sex

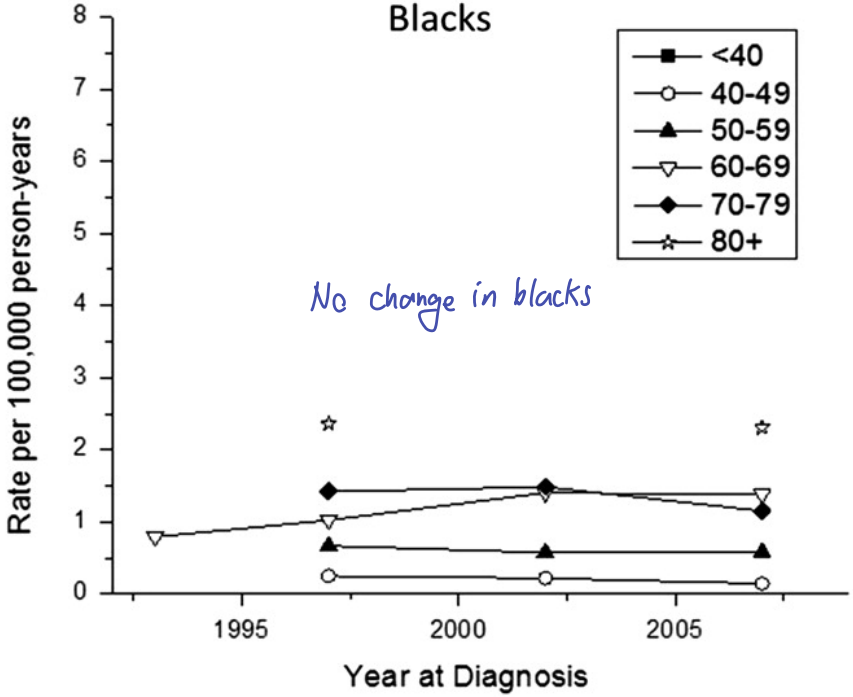




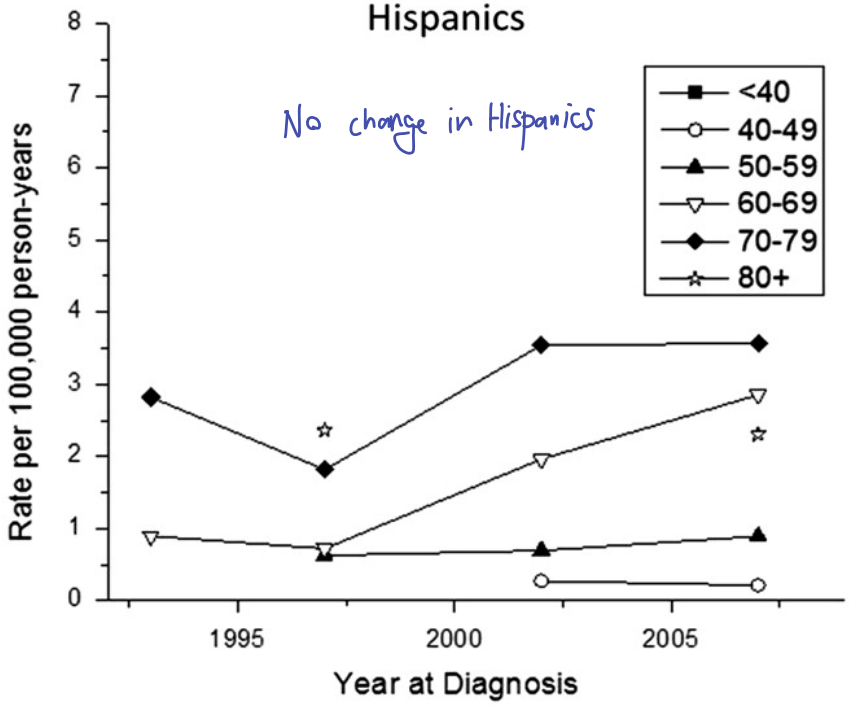
Stratified by race ethinicities



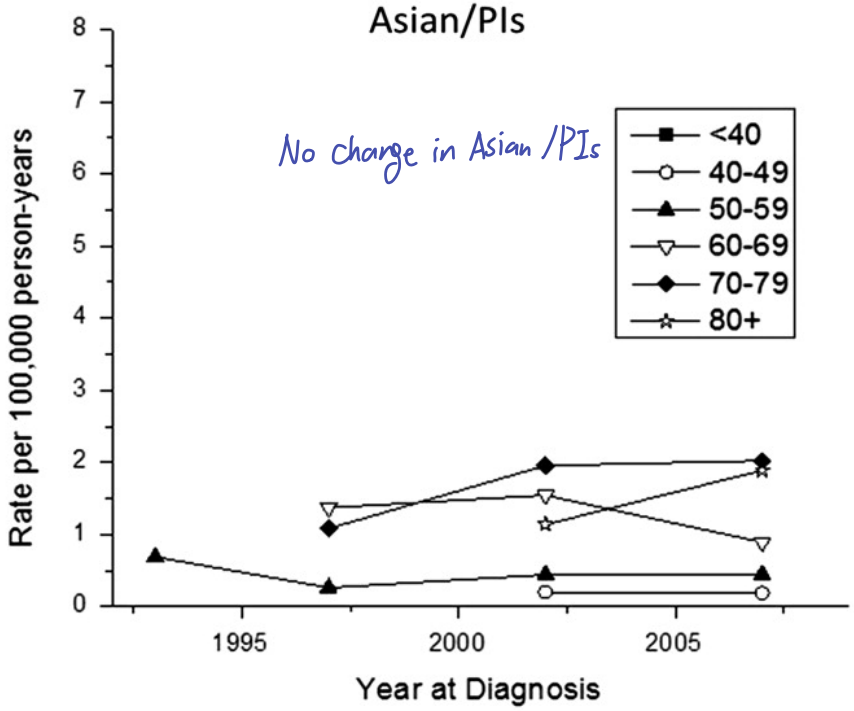
• Noticed ↑ in white people



No change in blacks



No change in Hispanics



No change in Asian/PIs

Conclusion: white males ↑↑↑.

Age, period, and cohort effects
(3 different ways time relates to disease)

Age as a Risk Factor

Age is associated with many demographic factors, health behaviors/lifestyle factors and with nearly all health conditions.

↑Age, ↑Exposures of behaviors, ↑chance for health effects

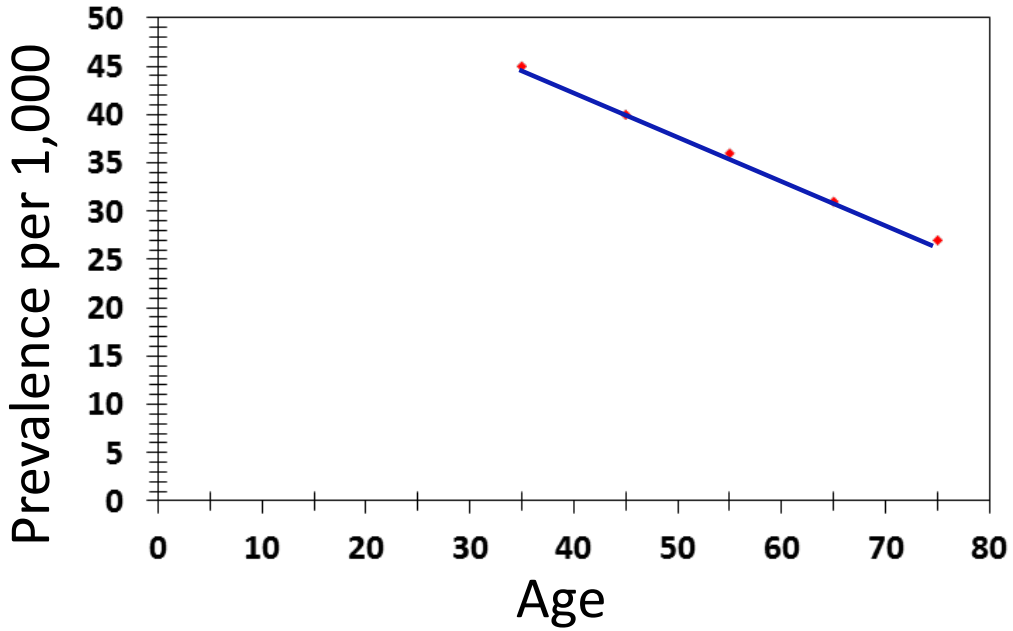
Demographics	**Behaviors**	**Health condition**
- ethnicity/race	- Smoking	- asthma
- education	- alcohol use	- cancer
- income	- screening	- depression
	- physical activity	- influenza
		- heart disease

Association of Age with Prevalence of Disease X (i.e., **Age Effect**)

- Question: Is there an association between age and the prevalence of disease X?
- **Age Effect:** Change in the rate according to age, irrespective of birth cohort and calendar time.

Age and Prevalence of Disease X in 2005

Age group	Midpoint	*2005 Prevalence*
30-39	35	45
40-49	45	40
50-59	55	36
60-69	65	31
70-79	75	27

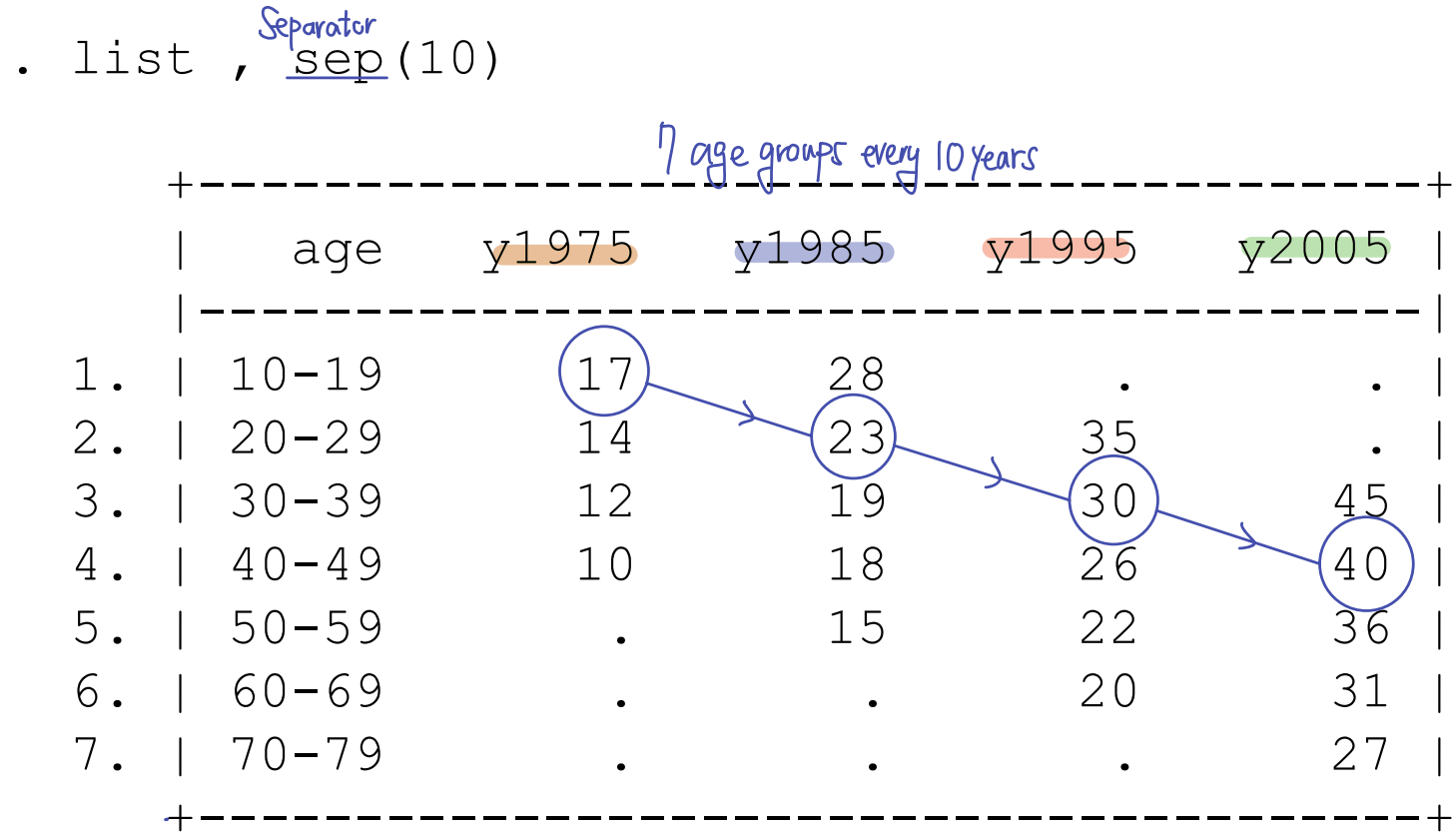


Question: Does this mean that the prevalence of disease X decreases with age? Yes.

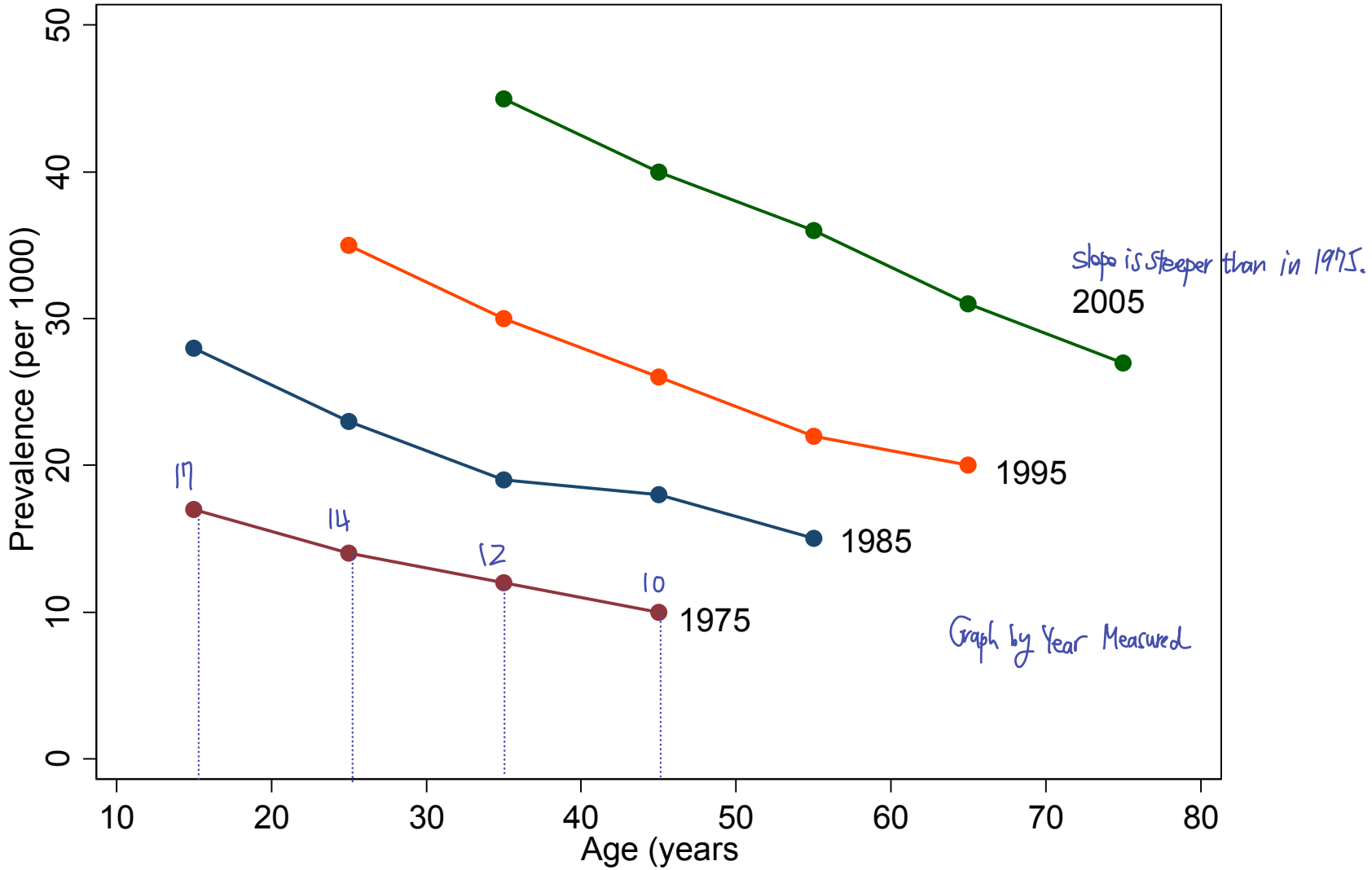
Hypothetical data from a series of cross-sectional studies of prevalence of disease X in a population, by age and survey date

```
. list , sep(10)

     +-------------------------------------------------+
     |    age    y1975    y1985    y1995    y2005     |
     |-------------------------------------------------|
  1. | 10-19        17       28        .        .     |
  2. | 20-29        14       23       35        .     |
  3. | 30-39        12       19       30       45     |
  4. | 40-49        10       18       26       40     |
  5. | 50-59         .       15       22       36     |
  6. | 60-69         .        .       20       31     |
  7. | 70-79         .        .        .       27     |
     +-------------------------------------------------+
```



Prevalence of disease X in a population by age and survey date



Birth Cohort Effect

Year of Birth

- **Age Effect:** Change in the rate according to age, irrespective of birth cohort and calendar time.

- **Cohort Effect:** Change in the rate according to year of birth, irrespective of age and calendar time.

"Same group"

Follow a group of people as they age

```
. list , sep(10)

     +-----------------------------------------+
     |   age   y1975   y1985   y1995   y2005  |
     |-----------------------------------------|
  1. | 10-19      17      28       .       .  |
  2. | 20-29      14      23      35       .  |
  3. | 30-39      12      19      30      45  |
  4. | 40-49      10      18      26      40  |
  5. | 50-59       .      15      22      36  |
  6. | 60-69       .       .      20      31  |
  7. | 70-79       .       .       .      27  |
     +-----------------------------------------+
```

7 age groups every 10 years

Birth Year

1966 - 1975 → 1970

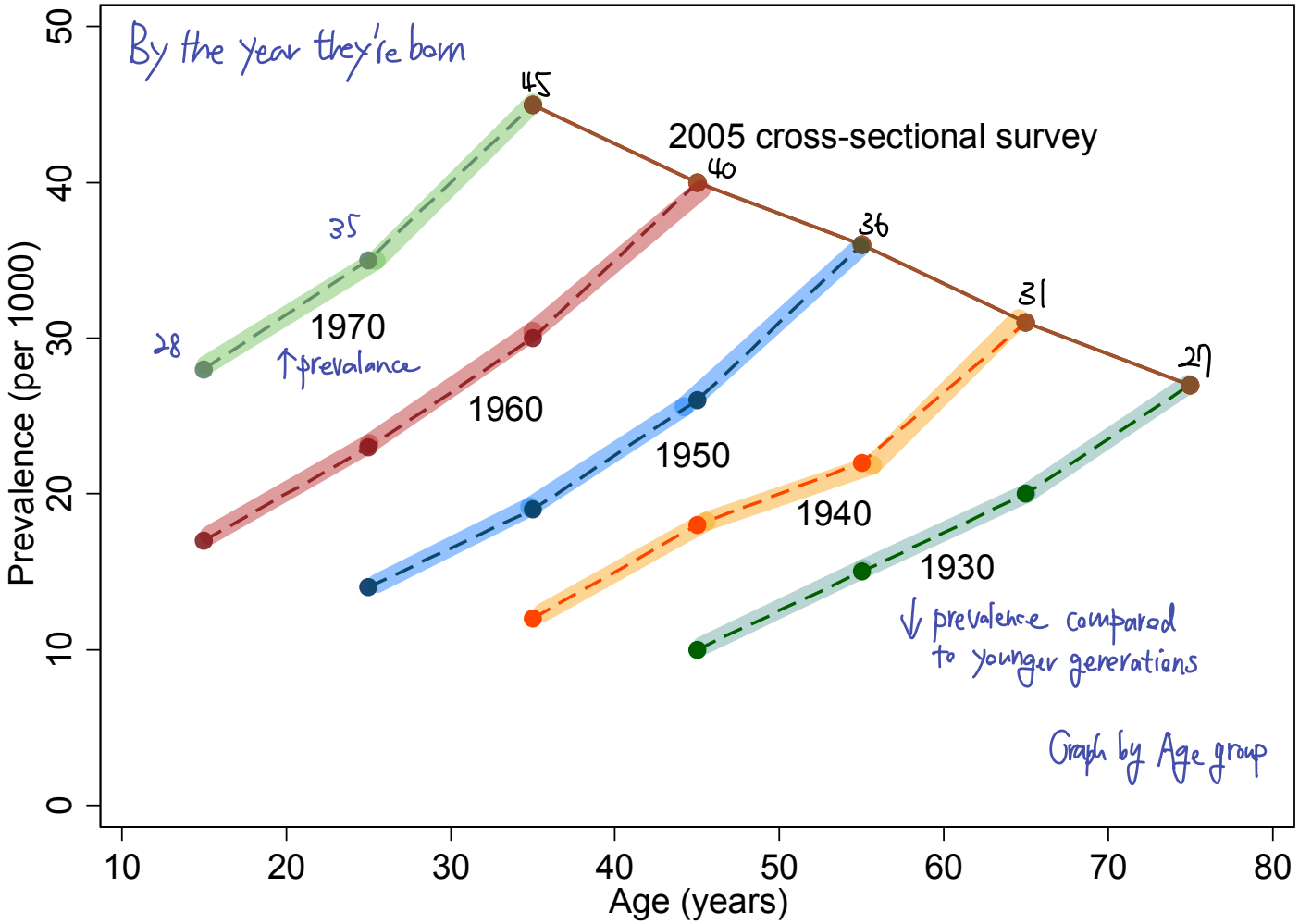
1956 - 1965 → 1960

1946 - 1955 → 1950

1936 - 1945 → 1940

1926 - 1935 → 1930

Plotting of the data in Fig 1.2 by birth cohort as they converge to the 2005 cross-sectional survey (solid line)



Does this mean that the prevalence rates of disease X decrease with age?

No for cohort effect. Each cohort of birth year actually showed ↑.

- **Age Effect:** rates vary by age in each cohort. Age
- **Cohort Effect:** rates are changing from cohort to cohort regardless of age. Birth Year

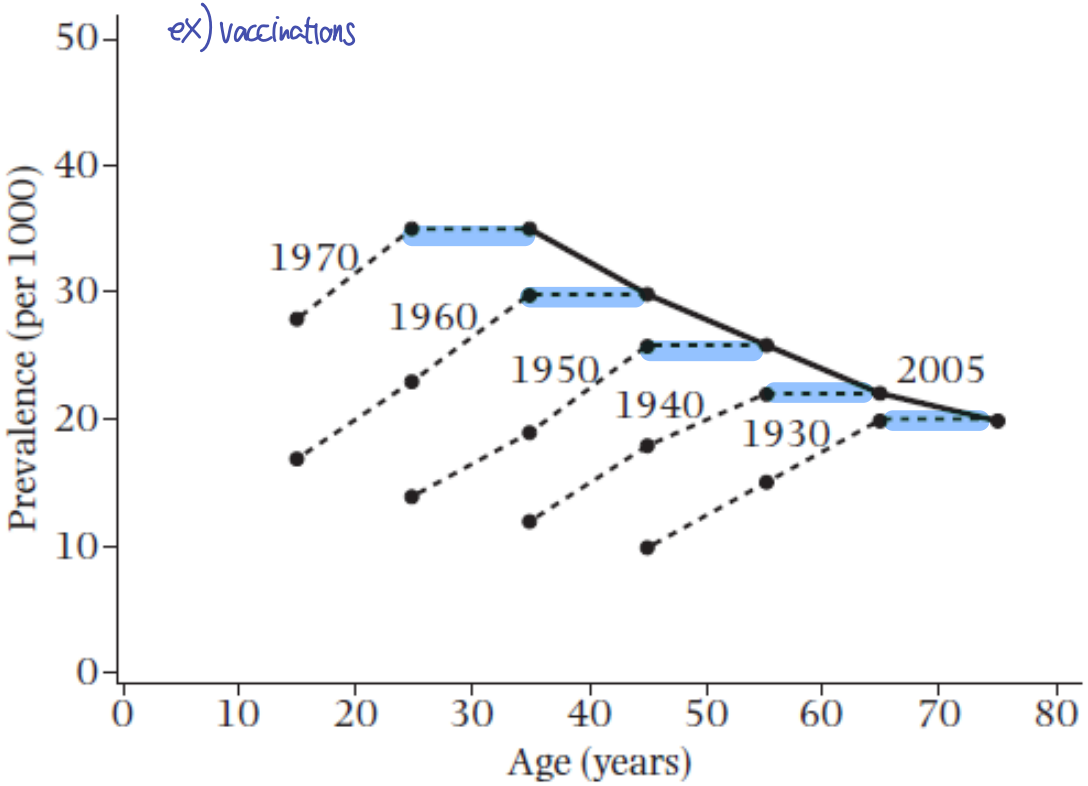
Period effect

- **Age Effect:** Change in the rate according to age, irrespective of birth cohort and calendar time.
- **Cohort Effect:** Change in the rate according to year of birth, irrespective of age and calendar time.

- **Period Effect:** Change in the rate affecting an entire population at some point in time, irrespective of birth cohort and age.

- **Period Effect:** Event in 1995 changes all cohorts in the corresponding ages.

FIGURE 1-5 Hypothetical example of period effect: an event happened in 1995 that affected all birth cohorts (1930 through 1970) in a similar way and slowed down the rate of increase with age. The solid line represents the observed cross-sectional age pattern in 2005.



Epidemiologic Study Designs

Descriptive: all are observational
hypothesis generating

> Surveillance
> Ecologic
> Cross-sectional

Analytic: test a hypothesis

> Observational
- Cohort
- Case-Control

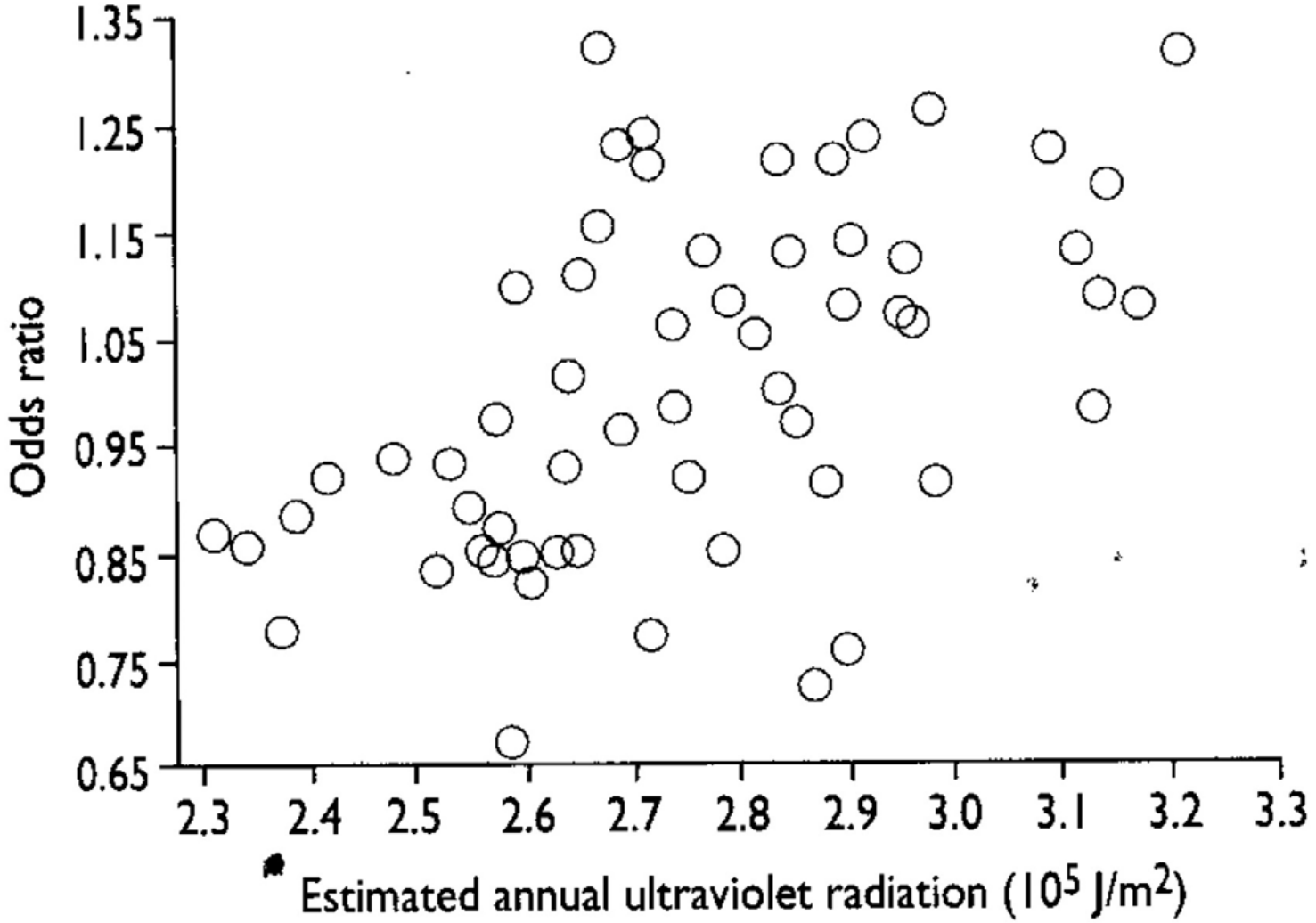
> Experimental
- Clinical Trials/Interventions
- Community Interventions

Ecologic Studies

- Unit of observation: NOT an individual but a group or population, such as a state, county, or country

- Proportion with exposure and proportion with disease are compared by group

Ecologic Study – Non-Hodgkin lymphoma and UV Radiation in 59 counties of England and Wales



Mortality of NHL

Bentham, G. BMJ 1996;312:1128-1131

bmj.com

Ecologic Studies – Major Concerns

- Ecologic fallacy
  - The incorrect conclusion that that occurs because the association observed at the aggregate level does not necessarily represent the association that exists at the individual level.

    Incorrectly draw conclusions of an individual (x) based on group level.
- Poor control of confounding
- Hypothesis generation

  generate hypothesis for us to use case-control + cohort to analyze the association.

Epidemiologic Study Designs

Descriptive: all are observational
hypothesis generating

> Surveillance
> Ecologic
> Cross-sectional

Analytic: test a hypothesis

> Observational
- Cohort
- Case-Control

> Experimental
- Clinical Trials/Interventions
- Community Interventions

Observational studies based on individuals

Cross-Sectional Studies

Exposure + Outcome are measured at the same time (cross-sectionally)

- Persons are enrolled AND THEN their exposure status and disease status are determined at a single point in time.
- Disease rates are compared for different levels of exposure
- Limitation:
- Whether disease or exposure occurred first cannot be determined in many cases
- Hypothesis generating

Wrong to assume gambling ↓ bronchitis:

ex)
- Not gamblers are too sick to gamble.
- Not gamblers are suggested by doctor to avoid smoking environment.

	# surveyed	# with bronchitis	Prevalence rate
Gamblers	100	2	20/1000
Not gamblers	100	5	50/1000

Epidemiologic Study Designs

Descriptive: all are observational
hypothesis generating

> Case report
> Case series
> Surveillance
> Ecologic
> Cross-sectional

Analytic: test a hypothesis

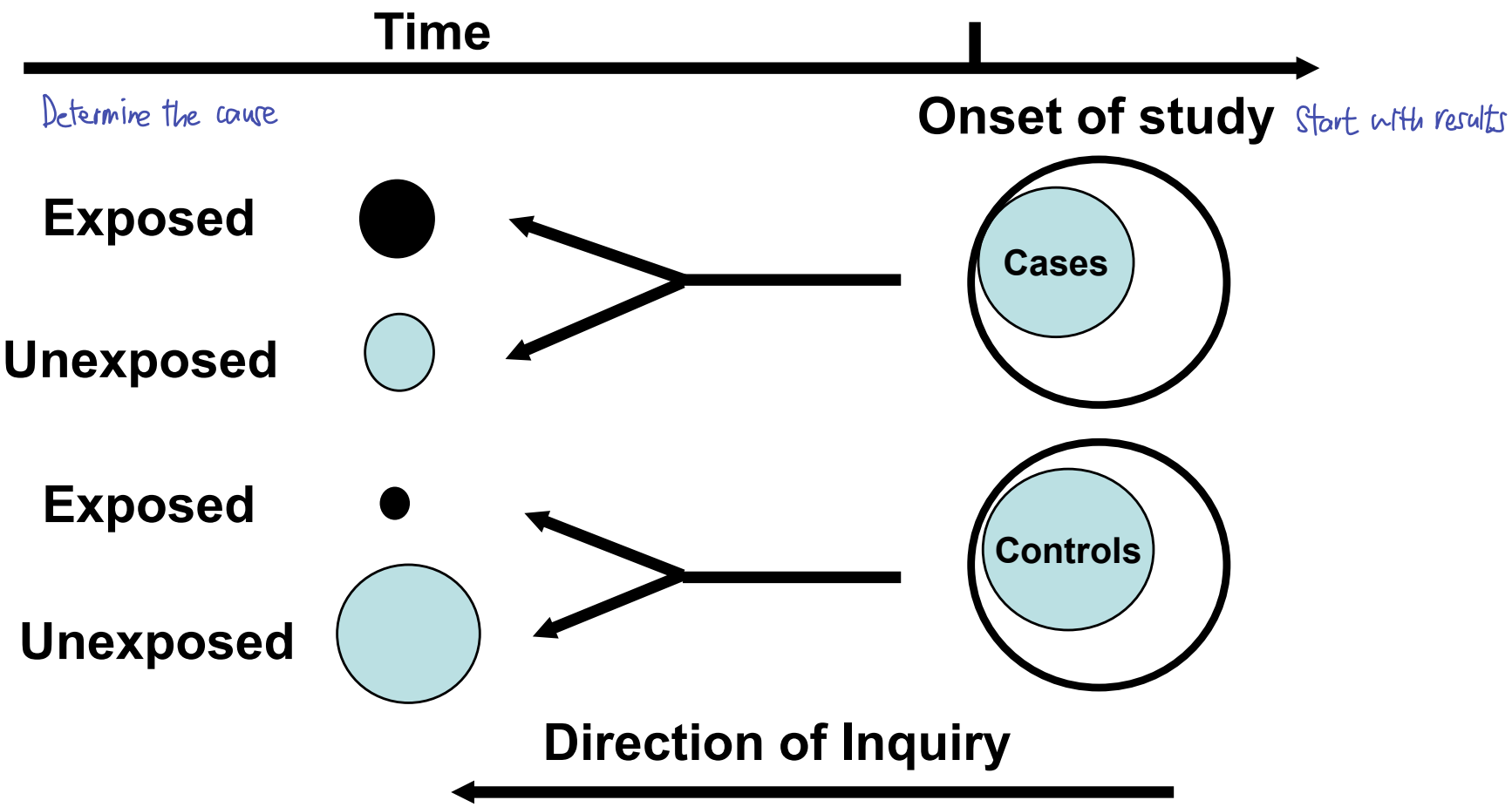
> Observational
- Cohort
- Case-Control

> Experimental
- Clinical Trials/Interventions
- Community Interventions

Case-Control Study Design

Disease - control study

- Persons with disease (case) and without disease (control) are enrolled
- Exposure status (prior to disease) of cases and controls is determined: by questionnaire, records, examination or testing
- Proportion with exposure compared for cases and controls



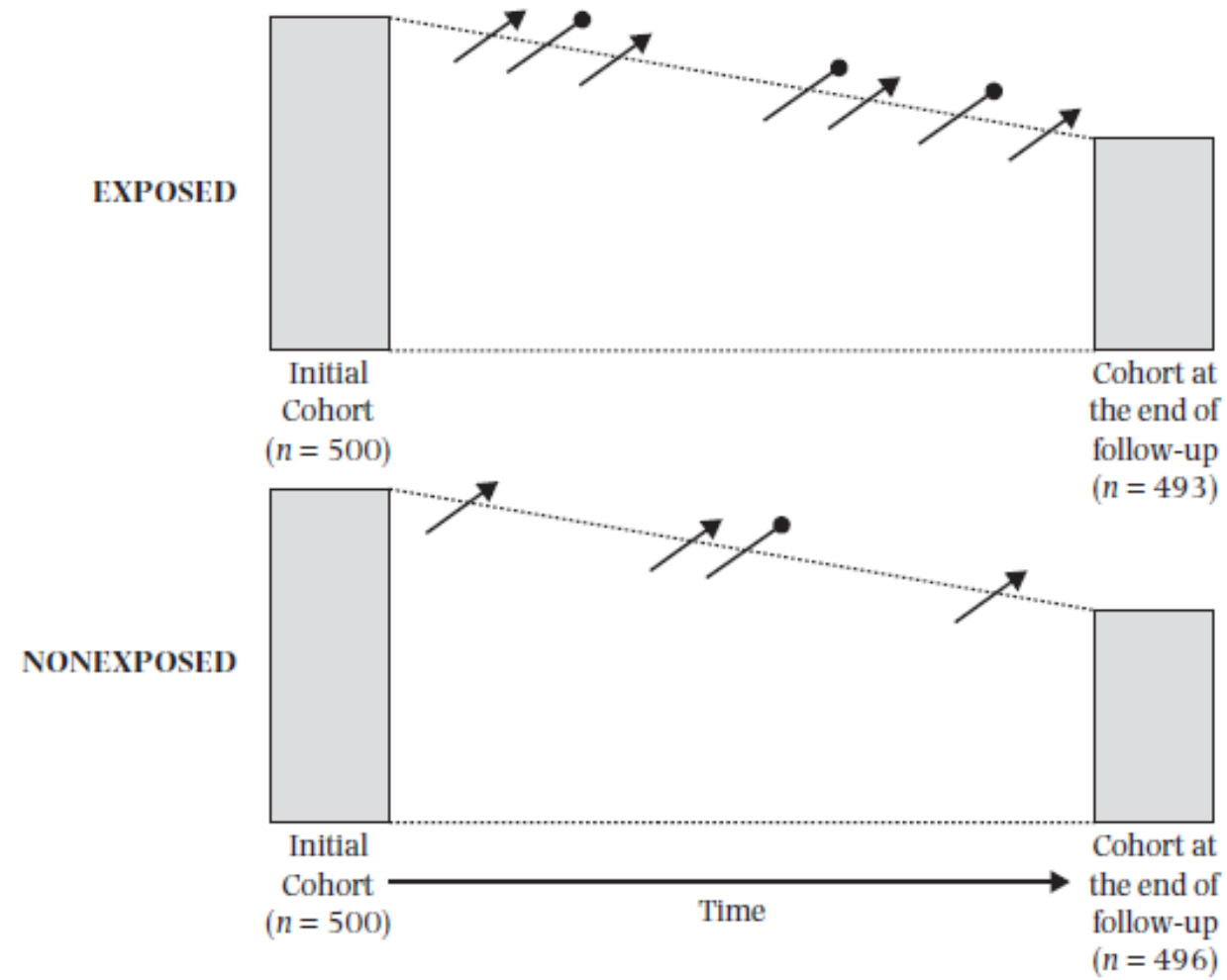
Cohort Studies

Start with cause
Determine the results

- Persons are enrolled on the basis of their exposure status, OR persons are enrolled and then their exposure status is determined
- They are followed over time for the occurrence of disease
- Disease rates are compared for different levels of exposure

How does exposure level influence rate of disease.

FIGURE 1-15 Same cohort study as in Figure 1-13, but the ascertainment of events and losses to follow-up is done separately among those exposed and nonexposed.



Within the participant serve as controls

☆Nested Case Control Studies

Allows for control of potential confounding variables through matching.

- Controls are a random sample of the individuals remaining in the cohort at the time each case occurs (i.e., incidence density sampling or risk-set sampling)
- Cases occurring later in the follow-up are eligible to be controls for earlier cases
- Equivalent to matching cases and controls on duration of follow-up which allows the use of straightforward statistical analysis techniques

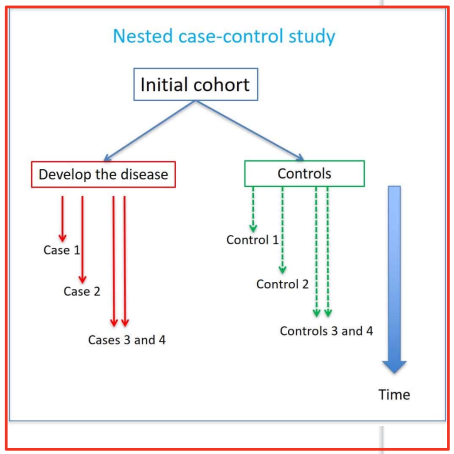
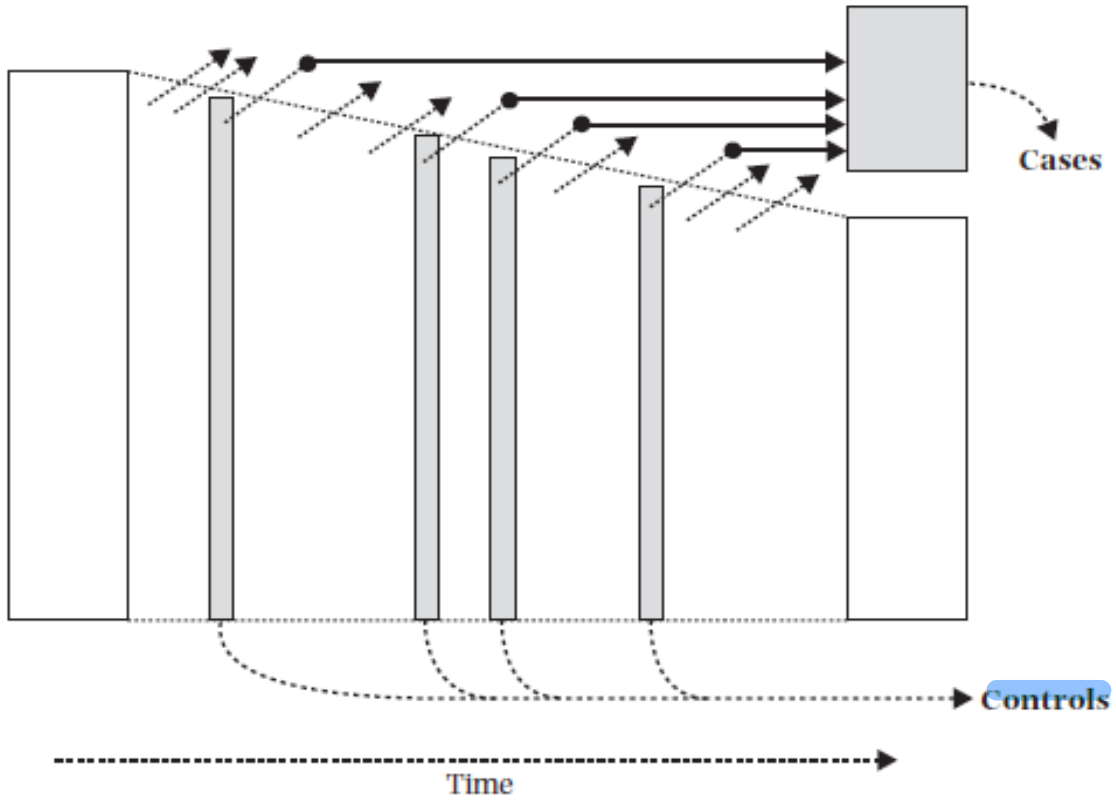


FIGURE 1-20 Nested case-control study in which the controls are selected at each time when a case occurs (incidence density sampling). Cases are represented by a dot connected to a horizontal arrow. Broken diagonal lines with arrows represent losses to follow-up.

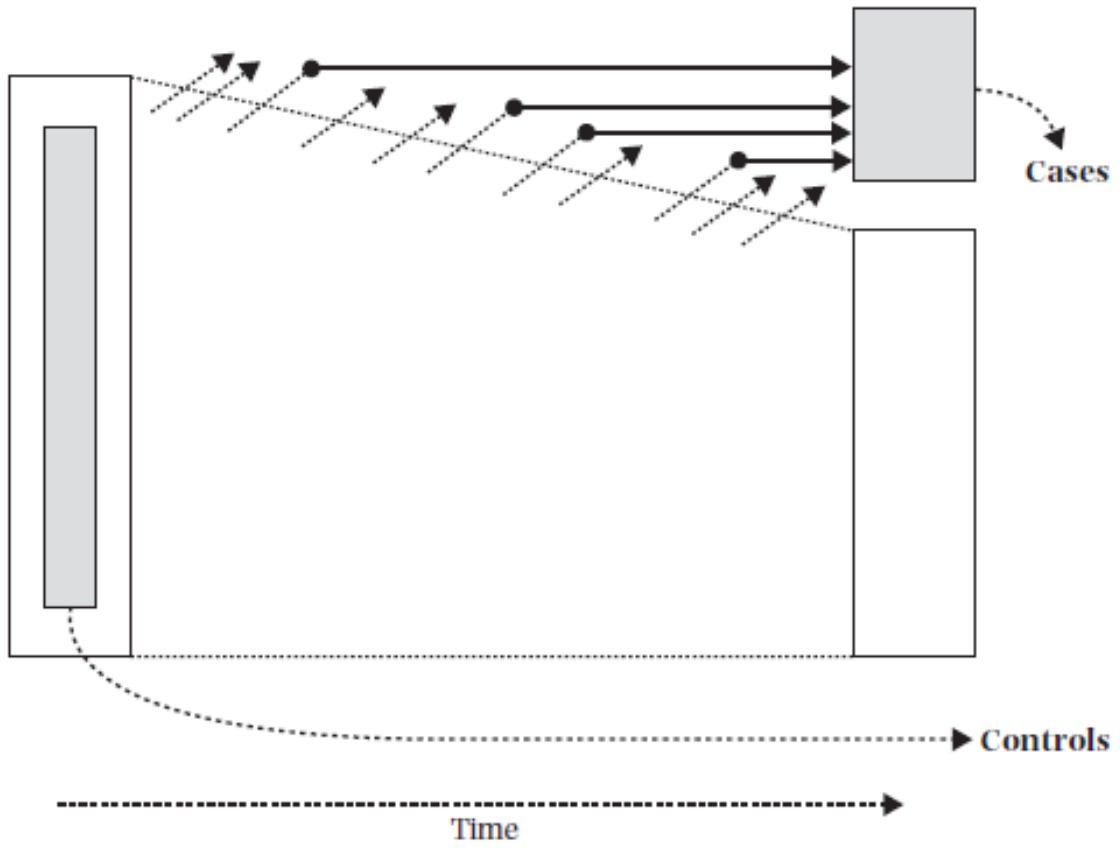


Limitation:
Reduced precision and power due to sampling of controls.

Case-Cohort Studies

- Controls are selected as a random sample of the total cohort at baseline
- The control group may include individuals who become cases during the follow-up
- Special statistical analysis methods needed due to the potential overlap between the case and the cohort random sample (control) groups

FIGURE 1-21 Case-control study in which the controls are selected from the baseline cohort (case-cohort study). Cases are represented by a dot connected to a horizontal arrow. Broken diagonal lines with arrows represent losses to follow-up.



When are these hybrid designs the best choice?

- In well-defined cohorts when additional (expensive or burdensome) information needs to be collected.
  - Laboratory determination in samples from specimen repository (e.g., serum bank).
  - Additional record abstraction (e.g., medical, occupational records).