Looking ahead…

- Intro to epidemiology; measures of occurrence
- Experimental, cohort, and case-control studies
- Cross-sectional and ecological studies
- Measures of association and potential impact
- Bias; confounding
- Interaction (effect modification); matching
- Measurement issues: validity and reliability
- Sampling methods
- Hypothesis testing and making inferences

What is epidemiology?

“The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to control of health problems.”

- Last’s “A Dictionary of Epidemiology”

The epidemiologic approach

- Observation / hypothesis
- Determine whether an association exists between “exposure” and “outcome”
- Evaluate the veracity of the observed association (or lack thereof)
- Is the association “causal”?

Epidemiology

- **Descriptive epidemiology**
  - Defines the frequency and distribution of health problems
  - Answers the questions: who, where, and when

- **Analytic epidemiology**
  - Analyzes the causes or determinants of health and disease
  - Answers the questions: why and how
### Descriptive epidemiology

- Organizing data related to health and health-related events by:
  - **Person (Who?)**
  - **Place (Where?)**
  - **Time (When?)**

### Person (Who?)
- Genetic factors
  - e.g., sex, race
- Biologic factors
  - e.g., age, nutritional status
- Behavioral factors
  - e.g., physical activity level, smoking status
- Socioeconomic factors
  - e.g., marital status, education level

### Place (Where?)
- Geographic distribution of disease
- Characteristics of where rates of disease are higher or lower

### Time (When?)
- Short-term fluctuations
  - Hours, days, weeks, months
- Period changes
  - Seasonal or cyclical
- Long-term changes
  - Years and decades

### Measures of occurrence
- **Incidence** (frequency of new “events”)
- **Prevalence** (frequency of existing “events”)

### Risk (or cumulative incidence)
- **Numerator**: # of events occurring in a specified population during a specified time period
- **Denominator**: # of persons in the specified population
- Restated…probability of developing the event during the specified time period
- Population must be at risk, and time period must be specified
- Values range from 0 to 1 (or 0% to 100%)
Can we think of some examples?

Prevalence

- **Numerator**: # of existing cases or events in a specified population at a specified point in time
- **Denominator**: Total population at the specified point in time
- Restated...proportion of the population who have the given condition or characteristic
- Values range from 0 to 1 (or 0% to 100%)

Incidence rate based on aggregate data

- Typically calculated for a geographic area or location (e.g., city, region, country)
- **Numerator**: # of events occurring in a specified population during a specified time period
- **Denominator**: average population size during the specified time period x specified time period
  - assumes an open (dynamic) but relatively stable population
  - for an annual rate: midyear population estimate x 1 year
- Example: incidence rate of gonorrhea in the US in 1989
  - 733,151 new cases reported
  - 246,552,000 population estimate (1989 midyear estimate)
  - incidence rate = 0.002974 cases per person-year
    - or 297.4 cases per 100,000 population (in 1989)

Relationship between incidence and prevalence

- **P** = Prevalence
- **I** = Incidence rate
- **D** = Duration of illness (on average)

\[ P \approx I \times D \]

- Assuming:
  - disease is in a steady state (incidence and duration have been relatively stable over time, and # of people entering the population = # of people leaving)
  - prevalence is low (e.g., <5%)

Uses of prevalence and incidence

- **Prevalence**
  - burden of an illness or attribute in a population
  - health care/public health planning
  - resource allocation
- **Incidence**
  - risk, probability of developing an outcome
  - etiologic studies
- Both are important for public health surveillance and health management