Descriptive Epidemiology

Part 1

Dr. H. Stockwell

Basic assumptions of epidemiology

- Human disease does **not** occur at random
- Causal and preventive factors can be identified through **systematic investigation** of **different populations** or subgroups of individuals in a population in **different places or times**

Epidemiology

- Divided into two major components:
  - Descriptive Epidemiology
  - Analytic Epidemiology (hypothesis testing)
- Both important to our understanding of disease
- Cannot ask relevant questions about disease etiology without a firm understanding of the descriptive epidemiology

Descriptive Epidemiology

- Designed to describe the existing distribution of variables without regard to causal associations
- Measure prevalence and incidence of disease/health
- Generate hypotheses for analytic studies
- **CANNOT TEST HYPOTHESES USING DESCRIPTIVE STUDIES**

Distribution of disease:

- **Person**: age, sex, race/ethnicity, SES, occupation, lifestyle
- **Place**: neighborhood, state, country, environment
- **Time**: date of exposure, date of diagnosis etc
**Race and Ethnicity in Epidemiologic Research**

- Often used variables in research – frequently used to assess the association of these variables on disease outcomes
- Biologically race is ill defined, poorly understood and may be of questionable validity
- Race has been described as an arbitrary system of visual classification (Fullilove, MT, 1998)
- DNA evidence indicates genetic diversity is a continuum with no clear breaks that delineate racial groups
- Since 2000 census individuals can self identify with more than one racial group

From Gordis L 3rd ed.

**Characteristics of person -- ex: gender and age**

Death rates (DR) per 100,000 population from coronary disease in the U.S., 1981, by age and gender

<table>
<thead>
<tr>
<th>Age</th>
<th>White men</th>
<th>White women</th>
<th>DR men/women</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-34</td>
<td>9.4</td>
<td>4.2</td>
<td>2.2</td>
</tr>
<tr>
<td>35-44</td>
<td>60.6</td>
<td>16.2</td>
<td>3.7</td>
</tr>
<tr>
<td>45-54</td>
<td>265.6</td>
<td>71.2</td>
<td>3.7</td>
</tr>
<tr>
<td>55-64</td>
<td>708.7</td>
<td>243.7</td>
<td>2.9</td>
</tr>
<tr>
<td>65-74</td>
<td>1669.9</td>
<td>769.4</td>
<td>2.2</td>
</tr>
<tr>
<td>75-84</td>
<td>3751.5</td>
<td>2359.0</td>
<td>1.6</td>
</tr>
<tr>
<td>85+</td>
<td>8596.0</td>
<td>7215.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

**Race and Ethnicity in Epidemiologic Research**

- Alternative approach is to use ethnicity
- Ethnicity is complex – may involve shared origins, culture, language
- What is the relationship to disease – does it increase our understanding of disease process, risk etc?

From Gordis L 3rd ed.

**Race and Ethnicity in Epidemiologic Research**

- Using race and/or ethnicity in our research may help us identify subgroups to which additional resources need to be directed.
- Some believe this further stigmatizes certain sub-groups
- If race is to be included in a study there should be a strong rationale, could other variables be better surrogates or is there a direct measure

From Gordis L 3rd ed.

**Place**

- Is there a geographic pattern?
**Characteristics of Place**

Across countries: latitude, temperature, amount of sunlight
Between cities or counties, areas: urban-rural and within a city

<table>
<thead>
<tr>
<th>G.B.</th>
<th>Japan</th>
<th>Nigeria</th>
<th>U.S.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver cancer</td>
<td>LOW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung cancer</td>
<td>HIGH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stomach</td>
<td>HIGH</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
<tr>
<td>Bladder</td>
<td>LOW</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
<tr>
<td>Colon</td>
<td>LOW</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
<tr>
<td>Prostate</td>
<td>LOW</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
<tr>
<td>Ovarian</td>
<td>LOW</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
<tr>
<td>Endometrial</td>
<td>LOW</td>
<td>LOW</td>
<td>HIGH</td>
</tr>
</tbody>
</table>

**Characteristics of Place:**

- Across countries: latitude, temperature, amount of sunlight
- Between cities or counties, areas: urban-rural and within a city

**Characteristics of Time Trends:**

- **Is there a temporal pattern? When is the disease occurring?**
  - Short term fluctuations in disease frequency – food borne outbreak
  - Cyclic patterns: annual increases in influenza in cold months
  - Secular trends: long term changes – over decades or more - heart disease

**Characteristics of Cyclical Time Trends**

- Periodic fluctuations on a seasonal basis/annual basis
  - Valuable mostly in investigation of acute diseases or those with a short latent period (period between exposure and disease onset/diagnosis)
  - Example: epidemiology of respiratory diseases/influenza – i.e., ease of transmission in winter months with increased crowding and human contact

**Characteristics of Secular Trends**

- Secular trends in chronic diseases may be caused by changes in:
  - Diagnostic techniques
  - Case finding
  - Accuracy in enumerating the population at risk
  - Age distribution of the population
  - Management of disease after diagnosis

Secular trends in chronic diseases may be caused by a change in the actual incidence of disease due to alterations in environmental, genetic or lifestyle factors.

Data from: National Linked Files of Live Births and Infant Deaths

Epidemic
- “the occurrence in a community or region of cases of an illness, specific health related behavior or other health related events clearly in excess of normal expectancy”
- Consider person, place and time

Disease Clusters
- “aggregation of relatively uncommon events or diseases in space and/or time in amount that are believed or perceived to greater than could be expected by chance”

Study Designs
- Studies are classified as **Descriptive** or **Analytic**
- **Descriptive** studies describe the situation – they **do not** test a hypothesis
- **Analytic Studies** test a hypothesis

Types of Descriptive Studies
- Case reports/case series: describe the experience of a patient or group of patients—may lead to a new hypothesis
- Correlational studies: measure characteristics in entire populations not individuals. May also be analytic and test a hypothesis
- Cross sectional surveys: exposure and disease measured at the same time in a group of individuals. May also be analytic and test a hypothesis

Case Report
Careful and detailed report by one or more clinicians of the clinical profile of a single patient

**Strengths:**
- Document unusual medical history/clinical features of disease
- Can provide clues in the identification of a new disease or adverse effects of exposures
Case Report

Limitations:
- There is nothing with which to compare the data - can’t determine whether something is unusual unless you know what’s usual
- Cannot be used to test for presence of a valid statistical association
- Since based on 1 person, presence of characteristic (risk factor - RF) may be coincidental

Case Series

Collections of individual case reports, usually within a fairly short period of time; description of clinical/epidemiologic characteristics of a number of patients with a given disease

Strengths:
- Better because it does not rely on a single case; shows better probability of a pattern
- Can examine the dose-response (D-R) relationship by examining the levels of exposure with the levels of disease severity

Limitations:
- No comparison group so cannot test for the presence of a valid statistical association

Correlational/Ecologic Studies

- Uses data from the entire population to compare disease frequencies between different groups during the same time period or same population at different points in time
- Example: per capita consumption of meat and colon cancer rates
- May be descriptive or analytic depending on whether testing a hypothesis

Ecologic Studies

Strengths:
- Cheap, quick
- May stimulate additional epidemiologic research
- May be the only design for uncovering assoc. at group level.
- Becoming popular again since there is GIS local level data

Weaknesses:
- Cannot link exposure with disease in individuals, therefore possibility of making an ecological fallacy.
- Sources of information may not be very accurate (Use average exposure levels rather than actual levels).
- Summary measures therefore imprecise Cannot control for confounding factors
- Cannot establish temporal sequence

Examples:
- Cigarette sales and mortality from CHD
- Death rates from breast cancer and dietary fat

Coronary heart disease mortality rate

[Graph showing coronary heart disease mortality rates vs. per capita cigarette sales]

Ecological Fallacy (also known as Aggregation Bias)

Patterns observed on the aggregate level are not observed at the individual level
- Cannot control for outside factors which may explain the association

Erroneous conclusions based on grouped data:
- The ecological fallacy refers to a bias that occurs when an association seen at the aggregate level does not represent the association seen at an individual level
- The association seen at the aggregate level is not true (biased association)

Many ecologic studies provide the basis for individual-level studies to be conducted - ecologic studies are often a good “first look”
Cross-Sectional Studies
- Exposure and disease outcome measured simultaneously
- Includes prevalent cases of disease (everyone with the disease at that point in time)
- No information on the temporal relationship between exposure and disease
- Good for variables that do not change (eye color, blood type etc) or good correlation between current and past practice - diet
- Both disease and exposure may have been the result of a third factor

Cross-sectional studies
- A snapshot (of a cohort) at one point in time
- Exposure and disease measured at same time
- Can compare (point) prevalence ratios or prevalence odds
- May be descriptive or analytic

Repeated Measures Studies
- Successive cross-sectional studies
- Repeated surveys of same population – not same individuals
- Detect overall time trends in a population

Person, Place and Time
- To learn more about the importance of place, please listen to the TED lecture