Concept of association and causation. Types of epidemiological studies I. Cross-sectional studies, cohort studies
Epidemiological studies

• Main functions are to:
  ✓ collect,
  ✓ analyze and
  ✓ utilize health-related information in order to improve population health.

• Planning epidemiological studies involves:
  ✓ professional (medical, epidemiological, ethical)
  ✓ administrative and
  ✓ economic considerations.
Epidemiological studies

• Most published studies are analytic or experimental

• These have the aim of discerning a cause-effect relationship between certain factors

• Different types of studies are able to provide different levels of evidence for a causal relationship

• In reality one-to-one cause-effect relationships are rare, we usually encounter „causal webs”

• Ascertainment of cause-effect relationships is one of the central and most difficult tasks of all scientific activities
Key components of epidemiological studies
Types of epidemiological studies

Epidemiological studies
  - Observational
    - Descriptive
      - Aggregate data
        - Ecological studies
    - Analytic
      - Individual data
        - Case-control studies (retrospective)
        - Cross-sectional studies (at a given time)
      - Cohort studies (prospective)
The typical course of epidemiological investigation

Considerable amount of anecdotal evidence, accidental findings... etc.} 

descriptive studies

Provisional hypothesis

Analytic studies

experimental studies

reviews of many experimental studies (i.e. meta-analyses)
Strength of evidence of studies

- Systematic review or meta-analysis of RCTs
  - Double-blind RCTs
  - Single-blind RCTs
- Randomized, controlled trials (RCTs)
- Non-randomized / uncontrolled experimental studies
  - „Regular“ cohort studies
  - Historical cohort studies
  - Case-control studies
  - Ecological studies
  - Cross-sectional studies
- Expert opinions, anecdotal reports
Descriptive studies

- Populational data
- Individual data (case, case-series)
  - Personal factors
    (age, gender, marital status)
  - Place
    (geographical and social environment)
  - Time
    (changes in a long or short period of time, seasonal changes)
Descriptive studies: the National Health Interview Survey 2003 (OLEF 2003)

Distribution of total amount of gratuities paid in 2003 between various health care sectors

- Hospital inpatient care (%): 42%
- Hospital outpatient care (%): 33%
- Specialist ambulant care (%): 25%
- GP services (%): 10%

Cross-sectional studies

These studies observe the
- exposition factor and the
- disease

at a same, given time on an individual level

These studies inform us about the frequency of the disease and the exposition factor at a given time so it estimates prevalence.
### Cross-sectional study

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Non-exposed</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

**Disease**

- $P = \frac{a}{a + b}$ vs. $P = \frac{b}{b + d}$
- $P = \frac{a}{a + c}$ vs. $P = \frac{c}{c + d}$
Cohort studies

Group of interest (e.g. smokers)

Follow over time

Comparison group (e.g. non-smokers)

Follow over time

Compare outcomes
# Cohort study

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exposed</strong></td>
<td>Yes: a</td>
</tr>
<tr>
<td><strong>Non-exposed</strong></td>
<td>Yes: c</td>
</tr>
</tbody>
</table>

**Study design**
Cannabis consumption and psychosis in a Dutch cohort study I.

- Van Os et. al. conducted the large-scale study between 1997 and 1999
- They selected 4045 non-psychotic individuals from the general population, and determined whether each individual used cannabis or not
- They then examined the incidence of psychosis in these subjects at 1 and 3 years

Cannabis consumption and psychosis in a Dutch cohort study II.

<table>
<thead>
<tr>
<th>Cannabis use</th>
<th>Psychosis</th>
<th>Healthy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>8</td>
<td>304</td>
<td>312</td>
</tr>
<tr>
<td>No</td>
<td>30</td>
<td>3622</td>
<td>3652</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>3926</td>
<td>3964</td>
</tr>
</tbody>
</table>

Measures of Risk I.

**Absolute risk**: incidence in a given population

*Risk of psychosis in exposed & unexposed?*

**Relative risk** (RR):

\[
\frac{\text{absolute risk in exposed } (R_{\text{exp}})}{\text{absolute risk in unexposed } (R_{\text{unexp}})}
\]

*Relative risk of psychosis in exposed compared to the risk in the unexposed?*

**Attributable risk** (AR): \( R_{\text{exp}} - R_{\text{unexp}} \)

*Risk of psychosis attributable to exposure?*
Attributable fractions

• **Attributable fraction in exposed**
describes the percentage of the incidence of the exposed group that occurs because of exposition

• **Attributable fraction in the whole population**
the proportion of the total risk of a disease in a population that can be attributed to exposure
Defining risk (formulas)

\[ I_{\text{exp}} = \frac{a}{a+b} \quad I_{\text{non-exp}} = \frac{c}{c+d} \]

Relative risk (RR) = \[ \frac{I_{\text{exp}}}{I_{\text{non-exp}}} \]

Attributable risk (AR) = \[ I_{\text{exp}} - I_{\text{non-exp}} \]

Attributable fraction in exposed(\%) = \[ \frac{I_{\text{exp}} - I_{\text{non-exp}}}{I_{\text{exp}}} \times 100 \]

Attributable fraction in population (\%) = \[ \frac{I_{\text{pop}} - I_{\text{non-exp}}}{I_{\text{pop}}} \times 100 \]
Some characteristics of cohort studies

• Exposure is measured prior to the onset of disease.
• The connection between an exposure and multiple outcome measures can be assessed simultaneously.
• Study design is prospective, but may be historic.
• Incidence can be measured directly.
• Relative and attributable risk can be calculated from incidence figures.
• Usually quite expensive and time-consuming.
• Studies typically require large efforts in organization and management, compliance of subjects is variable, many subjects may discontinue their participation.
Hill’s causal criteria

- **strength of association** (the stronger the more probable)
- **consistency** (over space, time, method, research group…)
- **dose - response relationship** (larger dose - larger effect)
- **chronological relationship** (cause before effect)
- **specificity** (one-to-one relationship)
- **biological plausibility** (is the relationship plausible at all?)
- **coherence** (does it fit with specific established „natural laws”)
- **analogy** (with similar systems of causation)
- **experimental evidence**
Guidelines for Assessing Causation (Bradford Hill, 1965)

- Temporal relationship correct
  - exposure must precede the outcome
Guidelines for Assessing Causation (Bradford Hill, 1965)

- **Strength of association**
  - the larger the relative effect, the more likely the causal role of the exposure.
    
    eg Smoking > 20 cigarettes/day – RR=20 of developing laryngeal cancer
  - Not all strong associations are causal
    eg Downs syndrome and birth rank
  - Weak associations do not rule out causality
    eg passive smoking and lung cancer (RR 1.4)
Guidelines for Assessing Causation (Bradford Hill, 1965)

- **Dose-response relationship**
  - if the risk increases with increasing dose of the exposure, the more likely the causal role of the exposure.

![Dose-response relationship diagram](image)

**FIGURE 4–1** An example of a dose-response relationship in epidemiology. The x-axis is the approximate “dose” of cigarettes per day, and the y-axis is the rate of deaths from lung cancer. (Source of data: Doll, R., and A. B. Hill. Lung cancer and other causes of death in relation to smoking. British Medical Journal 2:1071, 1956.)
Guidelines for Assessing Causation
(Bradford Hill, 1965)

• Replication of findings
  • if similar associations are found in different studies in
different populations, the more likely the causal role of
the factor.
   Eg Smoking and lung cancer
   • > 100 studies over last 30 years demonstrate increased
       risk
  • Lack of consistency does not rule out causality
    • blood transfusion not always a risk for HIV: virus must be
      present
  • Consistency may only be apparent when all relevant
details of cause are understood (seldom the case)
Guidelines for Assessing Causation (Bradford Hill, 1965)

• Association makes biological sense/plausibility
  • eg Histopathological effects of smoking on epithelium and
  • Depends on current knowledge
    eg John Snow and cholera epidemic in London (Vibrio cholerae was not yet discovered)
Guidelines for Assessing Causation (Bradford Hill, 1965)

• Consideration of alternate explanations
  • Takes into account the extent to which the researchers has considered alternative explanations for the outcome
    eg confounding
Guidelines for Assessing Causation (Bradford Hill, 1965)

• **Reversibility:**
  • Reduction or removal of the risk factor must reduce the risk of the outcome
  • Quitting smoking reduces risk
Guidelines for Assessing Causation
(Bradford Hill, 1965)

• Specificity of the association
  • Specific exposure is only related to one disease
  • Cigarette manufacturers use this to support their views
  • Not necessarily true where cigarette smoking is associated with lung cancer, heart disease, COPD, bladder cancer, etc
Guidelines for Assessing Causation (Bradford Hill, 1965)

• **Consistency with other knowledge**
  
  • Associations between the exposure and the outcome must be consistent with existing knowledge.
  
  • Variation in smoking between sexes → Difference in lung cancer incidence by sex