Epidemiology and Clinical Trials: A Methodological Approach

Christophe Lalanne

Fall 2011
Outline

What is epidemiology?  Common errors
Study design  What are clinical trials?
Advantages and drawbacks  Measuring treatment effect
Measures of association  The need for standardization
Causality, counterfactuals

“Life can only be understood backwards, but it must be lived forwards.” —Søren Kierkegaard (1813–1855)
The epidemiological way of thinking


The Disease Detectives detect patterns of disease in their line list

Interesting. Look at this: Everyone got ill 2-3 days before the school holidays.

Daniel said he got ill one day after the barbecue party. That was Saturday.
Epidemiology

One definition (Cates, 1982), among many others:

- A quantitative basic science built on a working knowledge of probability, statistics, and sound research methods;
- A method of causal reasoning based on developing and testing hypotheses pertaining to occurrence and prevention of morbidity and mortality;
- A tool for public health action to promote and protect the public’s health based on science, causal reasoning, and a dose of practical common sense.

Shortly, “Epidemiology is the study of the distribution and determinants of disease frequency in man.” (MacMahon and Pugh, 1970)
Various fields of research

For example,

- **Psychiatric epidemiology** is concerned with the study of prognosis and treatment of psychiatric disorders. Study on mental health related pathology does not always involve biological events or signals, thus emphasizing the need to build questionnaire or dedicated interview that might lead to some usable manifest variables. (Prince et al., 2003)

- **Genetic epidemiology** is the study of the joint action of genes and environmental factors in causing disease in human populations and their pattern of inheritance in families. (Thomas, 2004)
A common goal

The design and analysis of a study aims at maximizing the precision and validity of its findings. Avoiding bias or confounding effect is thus the primary goal of any epidemiologist, especially because most of the studies under consideration are observational per se. Strict control over experimental factors or exposure, like in a randomized controlled trial (RCT), are thus generally lacking.

The main objectives are generally one of explaining, monitoring, or forecasting.
Type of studies (Prince et al., 2003)

Observational / non-experimental
- Data from individuals
  - Descriptive
    - Cross-sectional survey
  - Analytic
    - Cohort study
- Data from groups
  - Descriptive
    - Case-control study
  - Analytic
    - Ecological study

Intervention / experimental
- Data from individuals
  - Randomized clinical trial
- Data from groups
  - Community trial
Examples (Ahrens et al., 2006)

The Doll and Hill’s study on smoking and lung cancer (1948–1952) was one of the first case-control study where 1,357 male and 108 female with lung cancer were recruited from several hospitals in London and matched with respect to age and sex to the same number of patients hospitalized for non-malignant conditions. A positive association between smoking and lung cancer was reported.

The Framingham Heart Study (1949–1999) included 5,127 participants free from coronary heart disease (CHD), aged 30 to 59 years, that were followed during 50 years to determine the rate of occurrence of new cases among persons free of disease at first observation. This study allowed to develop innovative statistical models and predictive risk measures. http://www.framinghamheartstudy.org/
During the Salk vaccine field trial (1954), nearly one million school children were randomly assigned to one of two groups, a vaccination group that received the active vaccine and a comparison group receiving placebo. A 50 percent reduction of the incidence of paralytic poliomyelitis was observed in the vaccination group as compared to the placebo group.
## GenomeWide Association Studies

(\texttt{http://pgc.unc.edu/})

### TABLE 6. Summary of Psychiatric GWAS Consortium Genomewide Association Study Samples and Characteristics of Studied Disorders$^a$

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Samples</th>
<th>Case Subjects</th>
<th>Comparison Subjects</th>
<th>Trios or Families</th>
<th>Prevalence Rate (%)</th>
<th>Heritability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD</td>
<td>6</td>
<td>1,418</td>
<td>0</td>
<td>2,443</td>
<td>4–12</td>
<td>70–80</td>
</tr>
<tr>
<td>Autism</td>
<td>6</td>
<td>652</td>
<td>6,000</td>
<td>4,661</td>
<td>0.3–0.6</td>
<td>90–100</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>10</td>
<td>7,075</td>
<td>10,559</td>
<td>0</td>
<td>0.3–1.5</td>
<td>73–93</td>
</tr>
<tr>
<td>Major depressive disorder</td>
<td>9</td>
<td>12,926</td>
<td>9,618</td>
<td>0</td>
<td>5–18</td>
<td>31–42$^b$</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>11</td>
<td>9,588</td>
<td>13,500</td>
<td>650</td>
<td>0.2–1.1</td>
<td>73–90</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>31,659</td>
<td>26,945</td>
<td>7,772</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Data shown are expected combined sample sizes for meta-analysis of genomewide association study data by the Psychiatric GWAS Consortium by the end of 2009. Data are reported for subjects of European-ancestry only; a small number of African American samples are also available for schizophrenia and bipolar disorder. The case subjects are all independent (although independence is tested using genotypes). For each disorder, comparison subjects used in more than one study are counted once; also, comparison subjects used for more than one disorder are counted once in the Total, which is therefore less than the sum of the rows for the disorders. The column for “Trios or Families” includes a sample of multiply-affected families for schizophrenia and trio or sib-pair families (with parents) for ADHD and autism. References for prevalence and heritability are as follows: ADHD (72–74), autism (75, 76), bipolar disorder (77, 78), major depressive disorder (79–82), and schizophrenia (83, 84).

$^b$ For major depression, higher estimates have been obtained in clinical samples (85) or using repeated interviews (81).
Advantages and drawbacks

- **Ecological studies**: Association on group level may be used for development of broad hypotheses; possible ‘ecological fallacy’ effect.

- **Cross-sectional surveys**: no direct evidence of causality but individual association may be used for development and specification of hypotheses; representative sampling frame, response rate, not applicable to rare disorders.

- **Case-control study**: Mainly retrospective, subjects defined by outcome, can handle multiple exposures, odds-ratio approximately equal to relative risk in a cohort study for rare disorders (prevalence < 10%), prone to several form of bias (selection, information, i.e. observer and recall).
• **Cohort study**: Longitudinal (prospective) design or historical cohort or population cohort, subjects designed by their exposure, not well-suited for rare disorders, can handle multiple outcomes/exposures, no information bias and an increased risk of disease among exposed will indicate a causal relationship.

• **Intervention study**: Allocating exposure to subjects and comparing the outcome of interest in different random groups, no information bias provided trials are double-blind; reduction of the incidence rate of the disease will confirm a causal relationship.

See also Ahrens et al., 2006, Table 1.2, p. 19.
### Recap’ on study design

<table>
<thead>
<tr>
<th></th>
<th>Cross-sectional</th>
<th>Case-control</th>
<th>Cohort</th>
<th>Ecological</th>
<th>Clinical trial</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subject selection</strong></td>
<td>Defined popula-</td>
<td>Caseness</td>
<td>Exposure</td>
<td>Aggregated data</td>
<td>Caseness</td>
</tr>
<tr>
<td></td>
<td>tion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Source of bias</strong></td>
<td>Selection, Non-</td>
<td>Selection, Information (Recall</td>
<td>Information (observer only), Loss</td>
<td>Selection of population, Ecological fallacy</td>
<td>Selection, Information (reduced by blinding)</td>
</tr>
<tr>
<td></td>
<td>response, Inform-</td>
<td>and observer)</td>
<td>to follow up (selection)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ation (Recall and observer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Probability of confounding</strong></td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>Very low if randomized</td>
</tr>
<tr>
<td><strong>Resources</strong></td>
<td>Quick and cheap</td>
<td>Relatively quick and cheap</td>
<td>Lengthy and expensive</td>
<td>Relatively quick and cheap</td>
<td>Relatively expensive</td>
</tr>
<tr>
<td><strong>Applications</strong></td>
<td>Planning ser-</td>
<td>Rare outcomes, Single outcomes, Multiple exposures</td>
<td>Rare exposure, Single exposures, Multiple outcomes</td>
<td>Rare outcomes, Rare exposures, Multiple exposures, Population exposures such as air pollution</td>
<td>Efficacy of new interventions, Effectiveness of new interventions, Hypothesis testing, Mechanisms</td>
</tr>
<tr>
<td></td>
<td>vices, Mapping secular and geographical trends, Identifying correlates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Measures of effect</strong></td>
<td>Prevalence</td>
<td>Odds ratio</td>
<td>Relative risk</td>
<td>Correlation or regression coefficient</td>
<td>Relative risk, odds ratio, difference between means</td>
</tr>
</tbody>
</table>
Correlation does not imply causation

Causation is by far the most difficult aspects of epidemiological research. Cohort and cross-sectional studies might both lead to confoundig effects. (Rothman and Greenland, 1998, Chap. 2)

Some prerequisites:

- The phenomena or variables in question must covary, as indicated, for example, by differences between experimental and control groups or by nonzero correlation between the two variables.

- The relationship must not be attributable to any other variable or set of variables, i.e., it must not be spurious, but must persist even when other variables are controlled, as indicated for example by successful randomization in an experimental design (no difference
between experimental and control groups prior to treatment) or by a nonzero partial correlation between two variables with other variable held constant.

- The supposed cause must precede or be simultaneous with the supposed effect in time, as indicated by the change in the cause occurring no later than the associated change in the effect.

While the first two criteria can easily be checked using a cross-sectional or time-ordered cross-sectional study, the latter can only be assessed with longitudinal data, except for biological or genetic characteristics for which temporal order can be assumed without longitudinal data. Of course, the situation becomes more complex in case of a non-recursive causal relationship.
Uncovering causal paths

**Independence**
- Factor A
- Factor B
- Outcome C

**Confounding**
- Factor A
- Factor B
- Outcome C

**Mediation**
- Factor B
- Factor A
- Outcome C

**Effect modification**
- Factor A
- Outcome C
- Factor B
The causality dilemma (Hill, 1965)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Strength</td>
<td>A strong effect size for an association</td>
<td>This reduces the chance of minor unmeasured confounding, but assumes that major confounding factors have been accounted for. Weak associations may also be causal.</td>
</tr>
<tr>
<td>2. Consistency</td>
<td>Repeated observations of an association in different populations</td>
<td>This assumes that all necessary causal factors are evenly distributed between populations. If a risk factor-outcome association were present only in men, would this imply non-causality?</td>
</tr>
<tr>
<td>3. Specificity</td>
<td>A risk factor leads to a single outcome</td>
<td>There is no reason why a risk factor should be associated with a single disorder (e.g. multiple disorders associated with alcohol misuse).</td>
</tr>
<tr>
<td>4. Temporality</td>
<td>The cause should precede the effect</td>
<td>A study should ideally demonstrate this. However the fact that one event follows another does not rule out the opposite direction of causation on other circumstances. For example, depression may cause physical ill-health but the opposite may also occur.</td>
</tr>
<tr>
<td>5. Biological</td>
<td>A 'dose-response' relationship</td>
<td>This assumes that the 'ceiling' of risk has not been reached. A single life-event may be sufficient to cause depression with no influence of further events. Of little use for cross-sectional associations since a 'dose-response' pattern of association would be predicted with either direction of causation.</td>
</tr>
<tr>
<td>gradient</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Plausibility</td>
<td>That the hypothesis is biologically plausible</td>
<td>Frequently a highly subjective judgement, given the volume of the biological literature. There are many historical examples of important findings rejected on the ground of implausibility at the time (e.g. Darwin's theory of the Origin of Species).</td>
</tr>
<tr>
<td>7. Coherence</td>
<td>That the interpretation does not conflict with the known biology</td>
<td>This depends heavily on the quality of the ancillary information. It also is not entirely consistent with the principle of hypothesis refutation.</td>
</tr>
<tr>
<td>8. Experimental</td>
<td>Evidence from interventional research</td>
<td>Intervenational research may not be ethical and/or feasible for many cause effect investigations. The intervention may not be discrete enough to infer causation.</td>
</tr>
<tr>
<td>evidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Analogy</td>
<td>Similar associations in other fields</td>
<td>A highly subjective judgement.</td>
</tr>
</tbody>
</table>
Common errors (Bhopal, 2009)

1. Failing to provide the context and definitions of study populations.
2. Insufficient attention to evaluation of error.
3. Not demonstrating comparisons are like-for-like.
4. Either overstatement or understatement of the case for causality.
5. Not providing both absolute and relative summary measures.
6. In intervention studies not demonstrating general health benefits.
7. Failure to utilise study data to benefit populations.
Recreating a virtual lab for health outcomes

I get frustrated trying to judge whether acne creams are having any effect. In the spirit of a controlled trial, I used one on just half my face for a few weeks.

It was cool seeing the effects so clearly, so I got some friends to try different treatments in an impromptu study.

Okay, you try the salicylic acid first. Wait, we should randomize the trials. Got a coin?

Okay, call it. Heads, she gets the— you!
Clinical trials

Clinical trials are clinical research studies to determine whether biomedical or behavioral interventions are safe, efficacious, and effective.

A more formal definition is available on NIH: http://www.nichd.nih.gov/health/clinicalresearch/.

Compared to observational studies, we aims at controlling all nuisance factors, much like in a designed experiment. Moreover, RCTs are registered, monitored, and their results should be reported according to standard guidelines.
Key concepts

• Use a prospective and comparative approach, with a control group and accounting for possible confusion factors.

• Use randomization and double-blinding to ensure that observed differences between groups come from treatment allocation only.

• No missing values and analysis by the intention-to-treat (ITT) approach. (No excluded patients, patients analyzed according to the randomization principle.)

“The effect of any treatment for a given patient is the difference between what happened to the patient as a result of giving the treatment and what would have happened had treatment been denied.” (Senn, 2007)
Types of clinical trial

Apart from the distinction between therapeutic trials (a new therapy is compared to a conventional one) and placebo-controlled trials (patients treated with a new treatment are compared to a group receiving a placebo), the pharmaceutical industry has established the following taxonomy: (Everitt and Wessely, 2004, Chap. 2)

- **Phase I trials:** clinical pharmacology and toxicity. Concerned with drug safety, not efficacy; usually performed on healthy, human volunteers ($n = 20 – 80$); characterization of the drug or treatment.
- **Phase II trials:** initial clinical investigation for treatment effect. Usually non randomized trials, with individual follow-up, to identify the patient population and study the effectiveness of the dosing regimen determined in Phase I.
• **Phase III trials:** full-scale evaluation of treatment. Comparison of the new drug with standard treatment in a large scale clinical trial; assessment of treatment efficacy and/or effectiveness.

• **Phase IV trials:** postmarketing surveillance. Monitoring for adverse effects, long-term studies of morbidity and mortality.
Common errors (Clark and Mulligan, 2011)

1. Failure to specify the inclusion and exclusion criteria.
2. Failure to determine and report the error of your measurement methods.
3. Failure to specify statistical assumptions made in the analysis.
4. Failure to perform sample size analysis before the study begins.
5. Failure to implement adequate bias control measures.
6. Failure to report missing data, dropped subjects and use of an intention to treat analysis.
7. Failure to perform and report power calculations.
Superiority, equivalence, non-inferiority (Lesaffre, 2008)
An observed association

Is it true in the study sample?
• Ensure data quality
  • Double check calculations and analyses

If it is true in the study sample, is it also true in the population?

Might it have arisen through sampling error (chance)?
• Consider confidence intervals
  • likely range of values in population
  • Consider p-value for probability of chance finding

Might it have arisen through error in the study design (bias)?
• Consider selection bias
  • Consider information bias
  • Consider likely influence of bias on observations

If it is probably true for the population, what can be inferred about cause and effect?
• Consider potential confounding factors
  • Consider residual confounding
  • Consider direction of causality
  • Consider mediating factors
  • Consider effect modification

What does this mean for me (as a researcher/clinician/public health worker)?

**CONSORT guidelines**
(http://www.consort-statement.org/):
A 25-item checklist and a participant flow diagram.

“Readers should not have to infer what was probably done, they should be told explicitly. Proper methodology should be used and be seen to have been used.” (Altman, 1996)
Participant disposition

1. Methods of recruitment and number obtained if multiple methods
2. Number excluded and reasons for exclusion
3. Number of candidates who refuse to enroll and reason(s) why (if after randomization from each treatment group)
4. Number of participants who withdraw from each treatment group and reasons why
5. Rate and nature of incidence of adherence with each treatment and reasons for nonadherence
6. Number lost to follow-up from each treatment group and reasons why
For the interested reader
Summary

• To ensure the validity of any findings, we must control for potential sources of bias and confounding.

• Statistical methodology comes into play when we design a study, not after data are observed.

• Although not discussed here, data collection, cleansing, pre- and post-processing are an important (and usually long) process.

• Data collected from groups can be used to formulate decision at the individual level.

• The choice of the outcome or phenotype to study is not always easy (e.g., genetic or psychiatric studies).
Bibliography

