

GRAU D'ESTADÍSTICA

ESTADÍSTICA MÈDICA

Francesc Miras

`francesc.miras@upc.edu`

Universitat de Barcelona – Universitat Politècnica de Catalunya



31st October, 2018

LITERATURE

- 92 Sta** Breslow, N. and N. Day (1980). *Statistical methods in cancer research*. International Agency for Research on Cancer.
- 92 Gor** Gordis, L. (2004). *Epidemiology*. W.B. Saunders.
- 92 Jew** Jewell, N. (2004). *Statistics for Epidemiology*. Chapman & Hall/ CRC.
- 62P Kah** Kahn, H. and C. Sempos (1989). *Statistical methods in epidemiology*. Oxford University Press.
- 62P Mcn** McNeil, D. (1996). *Epidemiological research methods*. Wiley.
- KL.9** Porta, M. (2008). *A Dictionary of Epidemiology*. 5th ed. Oxford University Press.
- 92 Rot** Rothman, K. (2002). *Epidemiology: an introduction*. Oxford University Press.
- KL.8** Rothman, K., S. Greenland, and T. Lash (2008). *Modern epidemiology*. Lippincott Williams & Wilkins.
- 92 Woo** Woodward, M. (1999). *Epidemiology study design and data analysis*. Chapman & Hall/ CRC.

TODAY'S PROGRAMME

INTRODUCTION

- Some notation and classification criteria

DESIGN OF EPIDEMIOLOGICAL STUDIES

- Classification of epidemiological studies
- Randomized epidemiological trials
- Cohort studies
- Case-control studies
- Population-based sampling: Cross-sectional studies
- Ecologic studies

WHAT IS EPIDEMIOLOGY?

DEFINITION

According to the old Greek:

$\epsilon\pi\iota$ (epi) = upon, among

$\delta\epsilon\mu\omicron\sigma$ (demos) = people, district

$\lambda\omicron\gamma\omicron\sigma$ (logos) = study

Rothman, Greenland, and Lash (2008):

Unfortunately, there seem to be more definitions of epidemiology than there are epidemiologists.

According to the Dictionary of Epidemiology (2008):

The study of the occurrence and distribution of health-related states or events in specified populations, including the study of the determinants influencing such states, (...)

CLASSIFICATION OF EPIDEMIOLOGICAL STUDIES

SOME NOTATION

One of the main goals of epidemiology is to analyze the relation between an exposure and a disease of interest, for example, working in a coal mine and lung cancer. We will discuss the advantages and disadvantages of the basic epidemiological study types presented below according to this goal. We will use the following notation.

D	Disease of interest present
\bar{D}	Disease of interest absent
E	Exposure of interest present
\bar{E}	Exposure of interest absent

CLASSIFICATION OF EPIDEM. STUDIES (CONT.)

CLASSIFICATION CRITERIA (HERNÁNDEZ ET AL. 2000)

- Assignment to study groups:
Random allocation vs. criterion-based assignment.
- Number of measures on study subjects:
Cross-sectional vs. longitudinal study?
- Prospective or retrospective study?
- Selection criterion:
None, exposure-based, or disease-based?
- Study units:
Individuals or whole populations?

DESIGN OF EPIDEMIOLOGICAL STUDIES

STUDY TYPES

We can basically distinguish the following study types:

- Randomized epidemiological trials
- Cohort studies
- Case-control studies
- Populations-based studies/ Cross-sectional studies
- Ecologic studies

Mixtures of these study types exist, for example, case-cohort studies.

DESIGN OF EPIDEMIOLOGICAL STUDIES (CONT.)

According to Olsen and Basso (2010),

“The study design should be dictated by its aim(s), if possible. The challenge is to find a design that most effectively and validly addresses the research hypotheses.”

They also point out that,

“An epidemiologic study is always a simplification of reality. (...) Since no study can specify all modifiers and confounders, this is generally acceptable and, possibly, the only way to proceed in many situations, but it is important to be aware of this.”

RANDOMIZED EPIDEMIOLOGICAL TRIALS

CHARACTERISTICS

- Sample from the population of interest.
- Random assignation of exposure.
- Longitudinal prospective study.
- Study units are disease-free individuals.

EXAMPLE

Smith-Sivertsen, T., E. Díaz, D. Pope, R. Lie, A. Díaz, J. McCracken, P. Bakke, B. Arana, K. Smith, N. Bruce (2009). Effect of Reducing Indoor Air Pollution on Women's Respiratory Symptoms and Lung Function: The RESPIRE Randomized Trial, Guatemala. *American Journal of Epidemiology* 170(2), 211–220.

RANDOMIZED EPIDEMIOLOGICAL TRIALS (CONT.)

ADVANTAGES

- Random exposure assignment minimizes possible bias.
- Prospective setting assures that exposure precedes disease.
- Risks $P(D|E)$ and $P(D|\bar{E})$ can be estimated and compared.
- High external validity.
- Several diseases may be studied.

DISADVANTAGES

- Possibly long study time and high costs.
- Possible lost-to-follow-up.
- Requires big samples in the case of rare diseases.
- In epidemiology, random exposure assignment is most often unfeasible or not ethic.

EXPOSURE-BASED SAMPLING: COHORT STUDIES

CHARACTERISTICS

- Selection criterion based on exposure of interest: samples are taken among exposed and non-exposed disease-free individuals, respectively.
- Longitudinal study
- Most often prospective
- Observational study

EXAMPLE

Muga, R., I. Ferrero, K. Langohr, P. García de Olalla, J. del Romero, *et al.* (2007). Changes in the incidence of tuberculosis in a cohort of HIV-seroconverters before and after the introduction of HAART. *AIDS* 21, 2521–2527.

COHORT STUDIES (CONT.)

ADVANTAGES

- Prospective setting assures that exposure precedes disease.
- Risks $P(D|E)$ and $P(D|\bar{E})$ can be estimated and compared.
- Allow to study rare exposures.
- Validity may be high (if groups are comparable).
- Several diseases may be studied.

DISADVANTAGES

- Possibly long study time and high costs.
- Possible lost-to-follow-up.
- Requires big samples in case of rare diseases.
- Exposure status may change over time.

DISEASE-BASED SAMPLING: CASE-CONTROL STUDIES

CHARACTERISTICS

- Selection criterion based on disease of interest: samples are taken among affected and disease-free individuals, respectively.
- Retrospective study

EXAMPLE

Cox, B. and M. Sneyd (2011). School Milk and Risk of Colorectal Cancer: A National Case-Control Study. *American Journal of Epidemiology* 173(4), 394–403.

CASE-CONTROL STUDIES (CONT.)

ADVANTAGES

- Ideal to study rare diseases or such with long latency.
- Several risk factors may be studied.
- Lower costs and less time consuming than cohort studies.

DISADVANTAGES

- Risks $P(D|E)$ and $P(D|\bar{E})$ cannot be estimated.
- It may not be clear whether exposure was prior to disease onset.
- Exposure may be difficult to determine exactly.
- Possible selection bias.

POPULATION-BASED SAMPLING: CROSS-SECTIONAL STUDIES

CHARACTERISTICS

- Sample from the population of interest
- One single measure.
- Presence/absence of both exposure and disease are determined.

EXAMPLE

Lewis, S., M. Antoniak, A. Venn, L. Davies, A. Goodwin, N. Salfield, J. Britto, and A. Fogarty (2005). Secondhand smoke, dietary fruit intake, road traffic exposures, and the prevalence of asthma: a cross-sectional study in young children. *American Journal of Epidemiology* 161(5), 406–411.

CROSS-SECTIONAL STUDIES (CONT.)

ADVANTAGES

- We can estimate $P(D)$, the disease prevalence, $P(E)$, and $P(D \cap E)$.
- Lower costs and less time consuming than cohort studies.
- May serve to generate hypothesis.

DISADVANTAGES

- Causal relations cannot be established, since it may not be clear whether E was prior to D .
- Incidence estimation is not possible.
- The investigator does not control proportion of individuals under exposure or with disease.

ECOLOGIC STUDIES

CHARACTERISTICS

- Uses aggregated data instead of individual data such as the mortality rates of several populations.

EXAMPLE

Green, C., L. Elliott, C. Beaudoin, C. Bernstein (2006). A Population-based Ecologic Study of Inflammatory Bowel Disease: Searching for Etiologic Clues. *American Journal of Epidemiology* 164(7), 615–623.

NOTE

Why are these studies called ecologic studies? See: Berry, B. (1970). Review of: Quantitative ecologic analysis in the social sciences by M. Dogan and S. Rokkan. *Geographical Review* 60 (1), 142–144.

ECOLOGIC STUDIES (CONT.)






ADVANTAGES

- Relatively easy to carry out.
- High statistical power due to high number of observations.
- Groups of study populations may be studied.

DISADVANTAGES

- Loss of information at individual level.
- Difficult to account for confounders.

REFERENCES

-  Hernández Avila, M., F. Garrido Latorre, and S. López Moreno (2000). Diseño de estudios epidemiológicos. *Salud pública de México* 42(2), 144–54.
-  Kreienbrock, L. and S. Schach (1995). *Epidemiologische Methoden*. Gustav Fischer Verlag Stuttgart·Jena.
-  Olsen, J. and O. Basso: Study designs. In: Olsen, J., R. Saracci, and D. Trichopoulos (2010): *Teaching Epidemiology*. 3rd ed. Oxford University Press.
-  Porta, M. (2008). *A Dictionary of Epidemiology*. 5th ed. Oxford University Press.
-  Rothman, K., S. Greenland, and T. Lash (2008). *Modern epidemiology*. Lippincott Williams & Wilkins.