Grau d'Estadística

Estadística Mèdica

Francesc Miras

francesc.miras@upc.edu

Universitat de Barcelona – Universitat Politècnica de Catalunya

14th November, 2018

Today's programme

MEASURES OF DISEASE - EXPOSURE ASSOCIATION

- THE RELATIVE RISK
  - Definition
  - Comments
  - The Risk Difference
- THE ODDS RATIO
  - Odds and Odds Ratio
  - Comments
  - Relation between Relative Risk and Odds Ratio
- THE ATTRIBUTABLE RISK
  - Population Attributable Risk
  - Exposed Attributable Risk

An introductory example

A hypothetical example

A cohort study on the possible effects of an exposure on a disease of interest has been carried out with 350 persons and after a certain time of follow-up, the following numbers of disease cases are observed:

	Disease		
Exposure	Yes	No	**Total**
Yes	45	105	150
No	40	160	200
Total	85	265	350

✍ Does the exposure increase the probability for the disease?

The Relative Risk

Definition

The **relative risk** (or **risk ratio**) is the ratio of the risk of disease (D) among exposed people (E) to the risk among unexposed people ($\bar{E}$):

$$\mathrm{RR} = \frac{\mathrm{P}(D|E)}{\mathrm{P}(D|\bar{E})}. \tag{1}$$

Note: The RR is the ratio of two cumulative incidences.

A hypothetical example (cont.)

In the previous example, the relative risk amounts to

$$\mathrm{RR} = \frac{45/150}{40/200} = \frac{0.3}{0.2} = 1.5.$$

That is, the risk of disease is 1.5 times higher among exposed people.

The Relative Risk (cont.)

Comments

- If RR > 1, there is a greater probability of D among exposed people. Hence, E is a (possible) **risk factor** for D.

  If RR < 1, E is a (possible) **protective factor** for D.

  RR $= 1$ indicates that there is no difference between both groups with respect to risk of disease. That is, D and E are independent.

- Since $\mathrm{P}(D|E) \leq 1$,

$$\mathrm{RR} = \frac{\mathrm{P}(D|E)}{\mathrm{P}(D|\bar{E})} \leq \frac{1}{\mathrm{P}(D|\bar{E})}.$$

  That is, the relative risk has an upper limit which is a function of the risk of disease among unexposed people.

- Do not only report the relative risk, but also $\mathrm{P}(D|E)$ and $\mathrm{P}(D|\bar{E})$ (Gigerenzer *et al.* 2011).

The Risk Difference

Definition

The **risk difference** (or **excess risk**) is the difference between the risk of disease (D) among exposed and unexposed people ($\bar{E}$):

$$\mathrm{RD} = \mathrm{P}(D|E) - \mathrm{P}(D|\bar{E}).$$

This measure *"...looks at the absolute, rather than the relative, difference in risk levels."* (Jewell 2004)

A real world example (Gigerenzer *et al.* 2011)

D: Thrombosis; E_1: Second-Generation Oral Contraceptives; E_2: Third-Generation Oral Contraceptives.

$$\mathrm{RR} = \frac{\mathrm{P}(D|E_2)}{\mathrm{P}(D|E_1)} = \frac{2/7000}{1/7000} = 2,$$

$$\mathrm{RD} = \mathrm{P}(D|E_2) - \mathrm{P}(D|E_1) = 2/7000 - 1/7000 = 0.000143.$$

The Relative Risk (cont.)

Comments (cont.)

- The RR can be estimated in cohort studies. To avoid bias, either follow-up should be similar for all subjects or lost-to-follow-up has to be unrelated to the risk of disease.
- The RR defined as in (1) is different to the ratio of two incidence rates (**incidence rate ratio**).
- In cross-sectional studies, the ratio of two prevalences is called the **prevalence relative risk (PRR)** or **prevalence ratio (PR)**.
- In case-control studies it is not possible to estimate $\mathrm{P}(D|E)$ and $\mathrm{P}(D|\bar{E})$ and since

$$\mathrm{RR} = \frac{\mathrm{P}(D|E)}{\mathrm{P}(D|\bar{E})} \neq \frac{\mathrm{P}(E|D)}{\mathrm{P}(E|\bar{D})},$$

  the relative risk cannot be estimated neither.

  Is there anything we can do to obtain $\widehat{\mathrm{RR}}$? $\rightsquigarrow$ Use the **odds ratio**!

The Odds Ratio

The Odds

The risk of disease D can also be expressed by means of the **odds**:

$$odds(D) = \frac{\mathrm{P}(D)}{1 - \mathrm{P}(D)}.$$

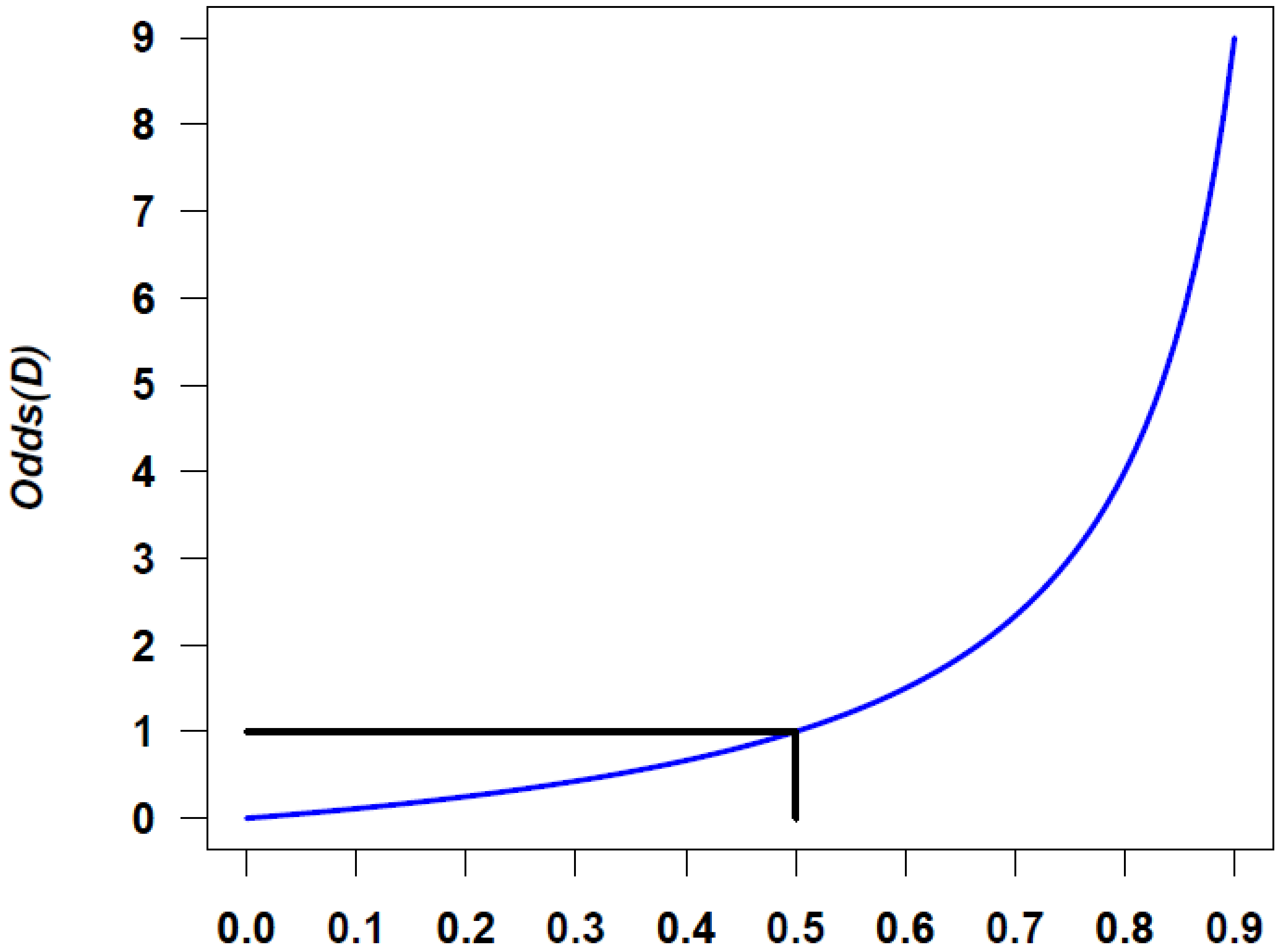
For example,

$$\left.\begin{array}{l} \mathrm{P}(D) = 0.2 \\ \mathrm{P}(D) = 0.5 \\ \mathrm{P}(D) = 0.6 \end{array}\right\} \Longrightarrow \left\{\begin{array}{lll} odds(D) = 0.2/0.8 & = 1/4 & (= 1:4) \\ odds(D) = 0.5/0.5 & = 1 & (= 1:1) \\ odds(D) = 0.6/0.4 & = 1.5 & (= 3:2) \end{array}\right.$$

If $\mathrm{P}(D) < 0.5$ ($\mathrm{P}(D) > 0.5$), then $odds(D) < 1$ ($odds(D) > 1$).

The odds is often used to describe the chance of winning a game.

The Odds Ratio (cont.)



The Odds Ratio (cont.)

Definition

The **odds ratio (OR)** is the ratio of the odds of disease among exposed people to the odds among unexposed people:

$$\mathrm{OR} = \frac{odds(D|E)}{odds(D|\bar{E})} = \frac{\mathrm{P}(D|E)/(1-\mathrm{P}(D|E))}{\mathrm{P}(D|\bar{E})/(1-\mathrm{P}(D|\bar{E}))}.$$

A hypothetical example (cont.)

In the previous example, the odds ratio amounts to:

$$\mathrm{OR} = \frac{45/150/105/150}{40/200/160/200} = \frac{45 \cdot 160}{105 \cdot 40} = 1.71.$$

That is, the odds of the disease is 1.71 times higher among exposed people than among unexposed people.

The Odds Ratio (cont.)

A second hypothetical example

In another cohort study, the following data were observed:

	Disease		
Exposure	Yes	No	**Total**
Yes	120	80	200
No	80	120	200
Total	200	200	400

✍ Calculate both the relative risk and the odds ratio and compare the results with the values of the first example. What do you observe?

The Odds Ratio (cont.)

Comments

- Since

$$\mathrm{OR} = \frac{\mathrm{P}(D|E)/(1-\mathrm{P}(D|E))}{\mathrm{P}(D|\bar{E})/(1-\mathrm{P}(D|\bar{E}))} = \mathrm{RR} \cdot \frac{1-\mathrm{P}(D|\bar{E})}{1-\mathrm{P}(D|E)}, \tag{2}$$

  the RR is always closer to 1 than the OR:

$$\left.\begin{array}{l} \mathrm{RR} < 1 \\ \mathrm{RR} = 1 \\ \mathrm{RR} > 1 \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \mathrm{OR} < \mathrm{RR} < 1 \\ \mathrm{OR} = \mathrm{RR} = 1 \\ \mathrm{OR} > \mathrm{RR} > 1 \end{array}\right.$$

- Contrary to the RR, the odds ratio has no upper limit, because

$$\lim_{\mathrm{P}(D|E)\to 1} \mathrm{OR} = \infty \quad (\text{in case } \mathrm{RR} > 1).$$

The Odds Ratio (cont.)

Comments (cont.)

- Contrary to the relative risk, the magnitude of the OR is difficult to interpret intuitively.
- Lee (1994):

  "(...), the odds ratio is incomprehensible."

- Holcomb et al. (2001):

  "There is no simple quantitative interpretation of the OR, except to the extent that it approximates the risk ratio."

The Odds Ratio (cont.)

Comments (cont.)

- Because of the relation in equation (2), RR $\approx$ OR as long as $\mathrm{P}(D|E)$ **and** $\mathrm{P}(D|\bar{E})$ are very small.
- In a case-control study, we can estimate $\mathrm{P}(E|D)$ and hence compute

$$\frac{\mathrm{P}(E|D)/(1-\mathrm{P}(E|D))}{\mathrm{P}(E|\bar{D})/(1-\mathrm{P}(E|\bar{D}))}. \tag{3}$$

  It can be shown that formula (3) is equivalent to the odds ratio. That is, the odds ratio can be estimated in case-control studies, too.
- Actually, case-control studies are often carried out when the disease of interest is rare; the RR can then be approximated by the OR.
- If the disease of interest is common, **interpret the odds ratio in terms of odds and not of risks!**

The Odds Ratio (cont.)

Comparison of the relative risk and the odds ratio

Table 1: Comparison of the RR and the OR (Source: Jewell (2004))

$P(D\|\bar{E})$	$P(D\|E)$	**RR**	**OR**
0.01	0.05	5	5.21
	0.1	10	11
0.05	0.1	2	2.11
	0.5	10	19
0.1	0.2	2	2.25
	0.5	5	9
0.2	0.4	2	2.67
	0.8	4	16

The Odds Ratio (cont.)

Comments (cont.)

- Another advantage of the OR deals with logistic regression: the model parameters can be interpreted in terms of the odds ratio, but not in terms of the relative risk.
- In cross-sectional studies, one can calculate the **prevalence odds ratio (POR)**.
- When comparing the risk of disease among more than two groups of exposure, one is chosen as the reference category and the OR (RR) is calculated for the others with respect to that one.
- In R, the RR and the OR can be calculated with the functions `riskratio` and `oddsratio` of the package `epitools`.

The Attributable Risk

Definition

The **attributable risk** (or **attributable fraction**) is defined as the proportion of disease cases attributable to a particular exposure.

1. The **Population Attributable Risk** is the proportion of disease cases in the population that is attributable to the exposure:

$$\mathrm{PAR} = \frac{N \cdot \mathrm{P}(D) - N \cdot \mathrm{P}(D|\bar{E})}{N \cdot \mathrm{P}(D)} = \frac{\mathrm{P}(D) - \mathrm{P}(D|\bar{E})}{\mathrm{P}(D)}.$$

With some transformations, we obtain:

$$\mathrm{PAR} = \frac{\mathrm{P}(E)(\mathrm{RR}-1)}{1+\mathrm{P}(E)(\mathrm{RR}-1)} = \mathrm{P}(E|D)(1 - \frac{1}{\mathrm{RR}}).$$

✍ Calculate the PAR in the previous examples.

The Attributable Risk (cont.)

Comments

- To estimate the PAR, the prevalence of exposure in the study population must be known or estimated.
- $P(D|\bar{E}) = 0 \Longrightarrow \text{PAR} = 1$; $P(D|\bar{E}) = P(D) \Longrightarrow \text{PAR} = 0$.
- Under certain assumptions, the PAR can be interpreted as the fraction of disease cases, which could be avoided, if the exposure was removed. But, have in mind that:
  - Exposure status might not be changeable.
  - This interpretation assumes a causal relation between E and D.
  - Removal of E might alter other risk factors.
  - The estimation of the PAR would need to be adjusted for all other possible causes of disease.
- If E is a protective factor for D, the PAR will be negative. However, in that case, PAR is not a meaningful measure.

The Attributable Risk (cont.)

Definition

2. The **Exposed Attributable Risk** is the proportion of disease cases among exposed people that is attributable to the exposure:

$$\mathrm{EAR} = \frac{\mathrm{P}(D|E) - \mathrm{P}(D|\bar{E})}{\mathrm{P}(D|E)},$$

which is equivalent to

$$\mathrm{EAR} = \frac{\mathrm{RR} - 1}{\mathrm{RR}} = 1 - \frac{1}{\mathrm{RR}}.$$

✍ Calculate the EAR in the previous examples.

The Attributable Risk (cont.)

Comments

- The EAR can be estimated in cohort studies.
- $\mathrm{P}(D|\bar{E}) = 0 \Longrightarrow \mathrm{EAR} = 1$; $\mathrm{P}(D|E) = \mathrm{P}(D|\bar{E}) \Longrightarrow \mathrm{EAR} = 0$.
- If the assumption of a low disease prevalence among both exposed and non-exposed people holds, the EAR can be estimated in case-control studies by means of:

$$\widehat{\mathrm{EAR}} \approx 1 - \frac{1}{\widehat{\mathrm{OR}}}.$$

- The expression in the numerator of the EAR is the **risk difference** or **excess risk**:

$$\mathrm{RD} = \mathrm{ER} = \mathrm{P}(D|E) - \mathrm{P}(D|\bar{E}).$$

- In R, both the PAR and EAR can be calculated with function `cs` of the `epiDisplay` package (Chongsuvivatwong 2015).

References

- Aragon, T. (2012). epitools: Epidemiology Tools. R package version 0.5-7. `http://CRAN.R-project.org/package=epitools`
- Chongsuvivatwong, V. (2015). epiDisplay: Epidemiological Data Display Package. R package version 3.2.2.0. `https://CRAN.R-project.org/package=epiDisplay`
- Gigerenzer, G., W. Gaissmaier, E. Kurz-Milcke, L.M. Schwartz, and S. Woloshin (2011). El significado de las estadísticas. *Mente y cerebro 50*, 62–69.
- Holcomb W., T. Chaiworapongsa, D. Luke, and K. Burgdorf (2001). An Odd Measure of Risk: Use and Misuse of the Odds Ratio. *Obstetrics and Gynecology 98*, 685–688.
- Jewell N. (2004). *Statistics for Epidemiology.* Chapman & Hall/ CRC.
- Kreienbrock, L. and S. Schach (1995). *Epidemiologische Methoden.* Gustav Fischer Verlag Stuttgart·Jena.

References

- McNutt, L.A., C. Wu, X. Xue, and J.P. Hafner (2003). Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *American Journal of Epidemiology 157(10)*, 940–943.
- Parzen, M., S. Lipsitz, J. Ibrahim, N. Klar (2002). An Estimate of the Odds Ratio That Always Exists. *Journal of Computational and Graphical Statistics, 11 (2)*, 420–436.
- Porta, M. (2008). *A Dictionary of Epidemiology.* 5^{th} ed. Oxford University Press.
- J. Lee (1994). Odds Ratio or Relative Risk for Cross-Sectional Data. *International Journal of Epidemiology 23(1)*, 201–203.
- Zhang, J. and K. Yu (1998). What's the relative risk? *Journal of the American Medical Association 280(19)*, 1690–1691.