Lecture 4

Hephaes Chuen Chau

2022-03-17

References: 1. Chapter 3 slides in MN

Introduction

- Measuring RR is usually the objective; a question in chapter 3 is how much difference does the point prevalence in cross-section study

3.2 Measuring Associations in a Cohort Study

- OR can also be calculated for a cohort study
- Note the algorithms to compute RR and OR for a two-by-two contingency table (p.6, slides)
- Note that in rare diseases (-ie, the incidence of such an event is low for both exposed sand non-exposed), OR is very close to RR (p.7 slides)
  - That is because, referring to the same contingency table in p.6, $a+b$ is very close to b as a is small (resp c is small in control)
- **Definition** of "built-in-bias":
  - Built-in-bias is defined to be

$$\frac{1-q_-}{1+q_+}$$

where q_+ is the probability of disease in the exposed group and q_- the probability of disease in the non-exposed group

- Values of OR
  - OR is directly computable from logistic regression, which is commonly applied
  - Reciprocal of OR wrt to event is the OR with respect to *non-event*
  - OR has the advantage of presenting the variation of risk in non-event, when the event of question is rare (p.12, slides)
- **Definition** of attributable risk AR_{exp} and (percentage attributable risks):
  - $\text{AR}_{\text{exp}} = q_+ - q_-$
  - $\%\text{AR}_{\text{exp}} = \left(\frac{q_+ - q_-}{q_+}\right) \times 100$
  - Alternative calculation method: $\%AR_{exp} = \left(\frac{RR-1.0}{RR}\right) \times 100$
  - Interpretation of :
    - The subjects of interest belong to the exposed
    - If the event of interest is disease, and exposure increases its risk: The interpretation is that is the percentage of the total risk of myocardial infarction among hypertensives that is attributable to hypertension
    - If the event of interest is disease, and exposure decreases its risk (eg, vaccines), we define a quantity with very similar form as :
      - Efficacy $= \left(\frac{q_{\text{cont}} - q_{\text{interv}}}{q_{\text{cont}}}\right) \times 100$

        - Alternative computation methods of efficacy: Efficacy $= 1 - \frac{1}{RR}$
- Population attributable risk (PAR)
  - To calculate PAR, we have to have q_{pop}, the incidence of the event of interest in the population of interest, then:
  - Pop AR $= q_{\text{pop}} - q_{-}$
  - % Pop AR $= \frac{(q_{\text{pop}} - q_{-})}{q_{\text{pop}}} \times 100$
  - How q_{pop} is calculated from exposure prevalence p_e:
    - $q_{\text{pop}} = [q_{+} \times p_e] + [q_{-} \times (1 - p_e)]$
    - It is possible that the %Pop AR$ can be calculated from p_e alone:
    - $\%\,\text{Pop AR} = \frac{p_e \times (\text{RR}-1)}{p_e \times (\text{RR}-1)+1} \times 10$
    - Interpretation: how much risk in having the event of interest in the population can be attributable to exposure
    - Utility: If one's research produces only p_e, the RR in the literature can be used to calculate the PAR (and vice versa).
    - From a public health policy angle, intervention is cost-effective when the event of interest is characterised by either a high RR or a high p_e, as these two quantities contribute to a higher (see the previous formula).
    - Important: in the above discussion, we have assumed throughout the RR is not produced in a cohort study affected by confounders
      - Example where this assumption is violated: age distribution is different between exposed and non-exposed group.
      - The standard advice is that *even after RR is adjusted for these confounders*, the above formula should NOT be used to calculate PAR.

3.3 Cross-sectional Studies: Point Prevalence Rate Ratio

- prevalence study can also be used to assess the association between exposure and disease; especially suitable if the exposure of interest is genetic elements (or in-born elements)
- with the assumption that the disease exists at a steady state - ie, number of individuals with the disease in a population is approximately constant, and the prevalence of the disease is very low, we can deduce that
  - Point Prevalence = Incidence $\times$ Duration $\times$(1$-$ Point Prevalence)
- In prevalence study the "association" between exposure and disease is represented by PRR
  - PRR $= \text{RR} \times \left(\frac{\text{Dur}_{+}}{\text{Dur}_{-}}\right) \times \left(\frac{1- \text{ Prev}_{+}}{1- \text{ Prev}_{-}}\right)$
  - Thus PRR can also represent RR if the two types of bias terms on the RHS contribute very little to PRR.

3.4 Measuring Associations in Case-control Studies

- the fact that RR should not be calculated in case-control study is illustrated by an example in which the control group is a sample of the non-diseased group in the cohort study
  - from the new data, the OR calculated is still the same as the OR calculated in RR
  - the hypothetical example also illustrates *the importance of ensuring random sampling in the selection of control - ie, we are not selecting more exposed individuals and less non-exposed individuals*
  - now observe that:
  1. *The rarity assumption:* in *rare* disease (disease with low incidence, *for the unexposed population*), OR calculated from case-control is the same as the OR calculated from cohort study, and it is close to RR

2. In non-rare disease, OR calculated from case-control is the same as the OR calculated from cohort study, and it is NOT close to RR

- the rarity assumption can be overcome if the control selected is a sample of the cohort (-ie, not necessarily only non-diseased individuals): the OR calculated will still approximate RR
  - this situation can be established by case-cohort study design. Case-cohort study also allows the population exposure prevalence to be estimated.