



CAUSATION AND ASSOCIATION AND CONFOUNDING

**DR DINESH KUMAR WALIA
ASSTT PROFESSOR
DEPARTMENT OF COMMUNITY MEDICINE
GOVERNMENT MEDICAL COLLEGE
CHANDIGARH**

Terminology

Risk ??

Risk Factor: Characteristic, condition or behaviour increasing the problem of developing a disease in a currently healthy individual.

i.e $P(D+ / E+) > P(D+ / E-)$

eg. for lung cancer: Smoking.

Risk Factor: env. / behavioural / social / psycho / genetic etc.
exposure to risk factor may be point / chronic.

Risk Assessment: Quantification / Extent of association / risk amount.

Risk Markers: Factors indirectly associated with the outcome.

eg. Neonatal Mortality:

Risk factors	:	Poor pre-natal care,
Casual factors	:	LBW
Marker	:	Low SES
Synergy	:	LBW & other complications.

Characteristics of Casual Factors:

- Consistency
- Strength
- Specificity
- Temporality
- Coherence

RISK FACTOR Vs CASUAL FACTOR

Casual Factor \equiv A risk factor which causes disease

Risk factor \neq casual factor (may be vice versa)

Why? The observed association between risk factor & disease may also be due to confounders / markers and synergistic factors.

Confounders??

Synergistic Factors?

Markers??

CONFOUNDING FACTORS

- Factor associated with exposure which independently of any such association is also risk factor for the disease.
- Factors distorting association between risk factor & disease by creating artificial relationships. A confounder is associated with both expo. and outcome, but unequally.

To eliminate the effect of confounding.

- Matching

To ensure comparability of cases and controls.

(a) Pair matching

(b) Group matching

- Stratification

Suspected etiologic factors should not be used as potential confounders for matching factors.

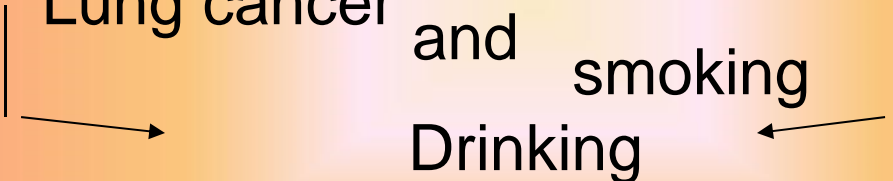
Examples of Confounding Factors

(1) Smoking & Lung cancer (age is confounder)


(2) Contra use (IUD) - Cervical cancer
(age & marital status)

(3) OC - BC (age)

Synergistic Factors: Synergy means magnification of effects arising from different exposures / risk factors s.t. effect of exposures working together is greater than the sum of their individual effects. ie. $P(AB) > P(A) + P(B)$.

eg. (1) A diagram for example (1) showing 'Lung cancer' and 'smoking' in a box, with arrows pointing down to 'Drinking'.

$R[\text{Smoke \& Drink}] > R(\text{Smoke}) + R(\text{Drink})$

(2) A diagram for example (2) showing 'MI' and 'OC use' in a box, with arrows pointing down to 'Smoking'.

Difference between confounding and synergy:

- In synergy, we are interested in joint effect or effect modification. A effect modifier (syn. Factor) relates essentially to differences between the measures of associate at different levels of modifier.
- In confounding, crude measure of associated does not reflect in each category necessarily, but the association is consistent.

ASSOCIATION AND CAUSATION

Types of Association:

- Spurious / unreal
- Indirect
- Direct / casual

Types of Causation

- One – to – One
- Multi-factorial

Criterion for judging Causality of Association

- Temporality
Whether suspected cause precede the observed effect.
- Strength of association
(RR or dose-response or duration response relationship)
- Specificity (One-One association, but a single cause may result in several outcomes)?
- Consistency
- Biological plausibility
- Coherence

Risk: Prob. / Prop. Of individuals suffering from particular disease out of total individuals.

$$\text{Risk: } P(D+) = \frac{\text{No. of diseased individuals}}{\text{Total no. of individuals}}$$

in cross sec. study, Risk = Prev. rate

Note: In case of cohort study (fixed follow-up)

$$\begin{aligned} \text{Risk} &\equiv \text{IR} \\ &= \frac{\text{No. of persons developing diseases during follow-up}}{\text{No. of persons disease free at start of study}} \end{aligned}$$

Probability and odds

- Probability
- Odds = No of times the event occurs/ No of times the event doesn't occur

= Probability that the event will occur divided by the Probability that the event will not occur.

$$= \frac{P \text{ (Occurrence of event)}}{P \text{ (No. occurrence of the event)}}$$

Odds of an event (say disease) is defined as the ratio:

$$\begin{aligned} &\equiv \frac{P(D+)}{P(D-)} \quad \text{or} \quad \frac{P(D+)}{1 - P(D+)} \\ &= \frac{a/N}{b/N} = \frac{a}{b} = \frac{\text{No. of occurrences}}{\text{No. of non-occurrence}} \end{aligned}$$

- Probability can be expressed as odds and odds can be expressed as probability.
- Odds = probability/1-probability= $P/1-P$
- $P = \text{odds} / 1 + \text{odds}$

Odds Ratio (OR) or Relative Odds (RO):

OR is used to compare absolute risks in two categories may be E+ and E- or D+ and D-

$$\text{OR} = \frac{(\text{Odds})_{\text{expo}}}{(\text{Odds})_{\text{unexpo}}}$$

In case of low risk / rare disease, OR gives an excellent approximation of relative risk (RR).

$$\text{OR} \approx \text{RR} \quad \hat{=} \quad \text{Estimated RR}$$

Relative Risk:

$$RR = \frac{\text{Risk among expo}}{\text{Risk among unexpo}}$$

In Cross-sectional study:

$$IR = \frac{\text{PR among expo}}{\text{PR among unexpo}}$$

In Cohort Study:

$$\equiv \frac{(\text{IR}) \text{ expo}}{(\text{IR}) \text{ unexpo}}$$

ATTRIBUTABLE RISK (AR)

Prop of diseased individuals attributable to a particular expo variable ie The amount of risk which can be attributed to a particular risk factor:

$$\begin{aligned}
 AR(1) &= \frac{P(D+/E+) - P(D+/E-)}{P(D+ / E+)} \\
 &= \frac{(IR)_{\text{expo}} - (IR)_{\text{unexpo}}}{(IR)} \\
 &\quad \text{(for follow-up studies)} \\
 &= 1 - \frac{1}{RR} = \frac{RR - 1}{RR}
 \end{aligned}$$

$$\begin{aligned} \text{AR (2)} &= P(E+/D+). \quad \text{AR (I)} \\ &= (\text{Expo rate among diseased}). \quad \text{AR (I)} \end{aligned}$$

$$\begin{aligned} \text{AR (3)} &= \frac{P(D+) - P(D+ / E-)}{P(D+)} \\ &= 1 - \frac{P(D+ / E-)}{P(D+)} \end{aligned}$$

where $P(D+) =$ Disease rate in the population ie PR.

Another form:

$$\text{AR (3)} = \frac{P(E+) (RR - I)}{1 + P(E+) (RR - I)}$$

- AR guides us in management of individual patients.
- AR gives the amount of risk that can be attributed to expo to a particular risk factor.

eg. Epilepsy & Delivery type

AR gives us the risk of epilepsy developing as a consequence of forceps delivery.

- Other syno terms of AR:

$$\begin{aligned} \text{Risk difference} &= P(D+/E+) - P(D+/E-) \\ &= (\text{Risk})_{\text{expo}} - (\text{Risk})_{\text{unexpo}}. \end{aligned}$$

≡ Excess Risk ≡ Rate Difference give crude expressions of AR.

POPULATION ATTRIBUTABLE RISK (PAR)

PAR is a measure of excess amount of risk of a particular disease in the population that can be solely attributed to the risk factor.

$$PAR = AR \cdot P(E^+) \text{ (in population)}$$

If an independent estimate $P(E^+)$ ie proportion of people exposed in the population is available, then PAR can be obtained.

TYPES OF BIASES / ERRORS IN EPIDEMIOLOGICAL STUDIES

