

Chapter 3: Epidemiologic Measures (Overview)

Epidemiologic measures are used to quantify the frequency of diseases in a population, measure the association between exposures and diseases, and address the potential impact of an intervention. Here's the outline of Chapter 3:

3.1 Measures of Disease Frequency¹

- Incidence Proportion (Risk, Cumulative Incidence)
- Incidence Rate (Incidence Density, Average Hazard, Person-Time Rate)
- Prevalence
- Odds²

3.2 Measures of Association between an exposure and outcome (disease)

- Absolute Measures of Associations (“Risk Difference”)
 - Risk Difference
 - Rate Difference
 - Prevalence Difference
- Relative Measure of Association (“Relative Risk”)
 - Incidence Proportion Ratio (Risk Ratio)
 - Incidence Rate Ratio (Rate Ratio)
 - Prevalence Ratio
 - Odds Ratio

3.3 Measures of Potential Impact

- Attributable Fraction in the Population
- Attributable Fraction in Exposed Cases
- Preventable Fraction

3.4 Rate Adjustment

- Direct adjustment
- Indirect adjustment (and the SMR)

¹ Measures of disease frequency are often called “rates” (even though only the “incidence density” is a true rate, mathematically speaking).

² Incidence proportion and prevalence “count data” may also be expressed as odds.

Measures of Disease Frequency (Formulas and Notes)

$$\text{Incidence Proportion} = \frac{\text{No. of onsets}}{\text{No. at risk at beginning of follow-up}}$$

$$\text{Incidence Rate} = \frac{\text{No. of onsets}}{\sum \text{person-time}}$$

- In a cohort, Σ person-time can be estimated by summing individual person-time.
- In an open population, Σ person-time \approx (average population size) \times (duration of follow-up).
- Actuarial adjustments may be needed when the disease outcome is not rare.
- Examples of vital rates:

$$\text{Crude birth rate (per } m) = \frac{\text{births}}{\text{mid-year population size}} \times m$$

$$\text{Crude mortality rate (per } m) = \frac{\text{deaths}}{\text{mid-year population size}} \times m$$

$$\text{Infant mortality rate (per } m) = \frac{\text{deaths } < 1 \text{ year of age}}{\text{live births}} \times m$$

$$\text{Prevalence Proportion} = \frac{\text{No. of cases}}{\text{No. of individuals in the study}}$$

- Also called prevalence proportion and point prevalence.
- The prevalence odds = prevalence / (1 – prevalence). When the disease is rare (or at least *not* too common), prevalence proportion \approx prevalence proportion.
- One may hear reference to a period prevalence. The period prevalence should be avoided because it confuses the concepts of incidence and prevalence.
- Prevalence is dependent of the incidence and average duration of disease according to this formula: Prevalence \approx (incidence rate) \times (average duration of illness)

Reporting

- To report a risk or rate as a uniconhort (i.e., number of individuals expected to produce one case), take its reciprocal and report the risk as 1 per “uniconhort.” For example, an incidence proportion of 0.0025 = 1 in $\frac{1}{0.0025}$ or “1 in 400.”
- To report a risk or rate “per m ” simply multiply by m . For example, an incidence proportion of 0.0010 = 0.0010 \times 10,000 = 10 per 10,000. Use a population multiplier that derives a nice whole number numerator (e.g.,
- Rates should be reported with 3 significant digit accuracy. (Intermediate calculations should use at least 4 significant digits.)

Measures of Association (Effect)

Notation and terminology: These concepts apply to incidence proportions, incidence rates, and prevalence proportions, all of which are loosely called “rates.” Let R_1 represent the rate or risk of disease in the exposed group and let R_0 represent the rate or risk of disease in the non-exposed group.

Absolute Measure of Association/Effect (Risk or Rate Difference) is captured in the rate or risk difference (RD)

$$RD = R_1 - R_0$$

Relative Measure of Association/Effect (Risk or Rate Ratio)

$$RR = \frac{R_1}{R_0}$$

Note: The relative effect of an exposure can also be captured by the SMR (see section on Rate Adjustment)

Measures of Potential Impact are used to quantify the effect of removing a hazardous exposure. The two main measures of potential impact are:

- The **attributable fraction among exposed cases** is $AF_e = \frac{R_1 - R_0}{R_1}$, or equivalently, $AF_e = \frac{RR - 1}{RR}$.
- The **attributable fraction in the population** is $AF_p = \frac{R - R_0}{R}$, or equivalent, $AF_p = AF_e \times p_c$ where p_c represents proportion of population cases that are exposed

2-by-2 Cross-Tabulation of Count Data (Incidence Proportions and Prevalences only)

	D+	D-	Total
E+ (Group 1)	A_1	B_1	N_1
E- (Group 0)	A_0	B_0	N_0
	M_1	M_0	N

For person-time data, delete cells B_1 and B_0 and let N_1 and N_0 represent the person-time in group 1 and group 0, respectively.

For **cohort** and **cross-sectional** data, the risk (count data) or rate (person-time data), $R_1 = \frac{A_1}{N_1}$ and

$R_0 = \frac{A_0}{N_0}$. For **case-control data**, ignore the formulas for risks and rate and use $OR = \frac{A_1 B_0}{A_0 B_1}$ as a

“Relative Risk” measure. For matched-pair data, see text. For matched tuple sets, use WinPEPI’s “PairsEtc” program E. ‘Yes-no’ variable: compare subjects with 2 or more controls.

Chapter 10: Screening for Disease

Reproducibility (Agreement)

		Rater B		
		+	-	
Rater A	+	<i>a</i>	<i>b</i>	<i>g</i> ₁
	-	<i>c</i>	<i>d</i>	<i>g</i> ₂
		<i>f</i> ₁	<i>f</i> ₂	<i>N</i>

$$p_{\text{obs}} = \frac{a + d}{N}$$

$$p_{\text{exp}} = \frac{f_1 g_1 + f_2 g_2}{N^2}$$

$$\kappa = \frac{p_{\text{obs}} - p_{\text{exp}}}{1 - p_{\text{exp}}}$$

Validity of an Ascertainment

	Disease +	Disease -	Total
Test +	TP	FP	TP + FP (those who test positive)
Test -	FN	TN	FN + TN (those who test negative)
Total	TP + FN (those with disease)	FP + TN (those w/out disease)	<i>N</i>

$$\begin{aligned} \text{SEN} &= (\text{TP}) / (\text{those with disease}) \\ &= (\text{TP}) / (\text{TP} + \text{FN}) \end{aligned}$$

$$[\text{note: TP} = (\text{SEN})(\text{TP} + \text{FN})]$$

$$\begin{aligned} \text{SPEC} &= (\text{TN}) / (\text{those without disease}) \\ &= (\text{TN}) / (\text{TN} + \text{FP}) \end{aligned}$$

$$[\text{note: TN} = (\text{SPEC})(\text{FP} + \text{TN})]$$

$$\begin{aligned} \text{PVP} &= (\text{TP}) / (\text{those who test positive}) \\ &= (\text{TP}) / (\text{TP} + \text{FP}) \end{aligned}$$

$$\begin{aligned} \text{PVN} &= (\text{TN}) / (\text{those who test negative}) \\ &= (\text{TN}) / (\text{TN} + \text{FN}) \end{aligned}$$

$$\text{True prevalence} = (\text{TP} + \text{FN}) / N \quad [\text{also known as } \textit{prior probability}]$$

Bayesian equivalents for determining predictive values based on prior probabilities and test parameters are available in the text.



TEN COMMANDMENTS FOR DEALING WITH CONFOUNDING



Source: EPIB-601 McGill University, Montreal, Canada madhukar.pai@mcgill.ca,
<http://www.teachepi.org/documents/courses/Ten%20Commandments%20for%20Dealing%20with%20Confounding.pdf>

- I. Always worry about confounding in your research, especially at the design/protocol stage. Try to use design elements (e.g. randomization) that will help reduce potential confounding.
- II. Prior to the study, review the literature and consider the underlying causal mechanisms (e.g. draw causal diagrams such as directed acyclic graphs [DAGs]). Then make sure you collect data on all potential confounders; otherwise you will not be able to adjust for them in your analyses.
- III. Know your field or collaborate with an expert who does! Subject-matter knowledge is important to recognize (e.g. draw causal diagrams) and adjust for confounding.
- IV. Use a priori and data-based methods to check if the potential confounders are indeed confounders that should be adjusted for.
- V. Use stratified analyses and multivariable methods to handle confounding at the analysis stage. Choose the multivariate model that best suits the type of data (e.g. dichotomous vs. continuous) you collected and the design you employed (e.g. case-control vs. cohort).
- VI. Do not adjust for covariates that may be intermediate causes (on the causal pathway between the exposure and disease). Do not adjust for covariates that may not be genuine confounders. And beware of time-varying covariates that will need special approaches.
- VII. Use matching with great caution. Use analytic methods that are appropriate for the design used; for example, if matching was done, use methods that take matching into account (e.g. conditional logistic regression, matched pairs analyses).
- VIII. Always consider effect measure modification, but perform and interpret subgroup analyses with caution. The subgroup analysis should be one of a small number of hypotheses tested, and the hypothesis should precede rather than follow the analysis (i.e. subgroups must be pre-specified).
- IX. Always remember that adjustment for confounding can be inadequate due to residual confounding because of unmeasured confounders, misclassification of confounders, and inadequate adjustment procedures (e.g. model misspecification, categorization of continuous covariates).
- X. If conventional methods prove to be inadequate, consider using newer approaches such as propensity scores, matched sampling, instrumental variables and marginal structural models. However, make sure you work with statisticians who understand these new methods (not many do).

When all else fails, pray! If prayer fails, consider changing professions!!