Chapter 6: Incidence and Prevalence

Incidence Proportion (IP) = $\frac{\text{onsets}}{\text{no. at risk}}$

1. Synonyms: average risk, cumulative incidence
2. Cohorts only, length of follow-up should be specified

Incidence Rate (IR) = $\frac{\text{onsets}}{\sum \text{person - time}}$

1. Synonyms: incidence density, person-time rate
2. Person-time in cohorts: tally individual person-time
3. Person-time in open population ≈ (average population size) × (time)
4. IR ≈ one-year IP when risk ≤ 5%
5. IR = inverse of survival time (if steady state population or cohort with full follow-up)
6. Examples of rates in open populations

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

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

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

Prevalence (P) = $\frac{\text{cases}}{\text{No. in population}}$

1. Depends on inflow and outflow (Fig. 6.9)
2. Prevalence ≈ (incidence rate) × (average duration)

Incidences and prevalence may be reported with a population multiplier, i.e., "per *m* people". To convert a rate or proportion to "per *m* people," simply multiplying by *m* (e.g., 0.008770 per year 0.008770 per year × 100,000 = 877 per 100,000 person-years

Chapter 8: Measures of Association and Potential Impact

Let R_1 represent the rate or risk in the exposed group and R_0 represent the rate or risk in the non-exposed group.

The **Risk Difference** $RD = R_1 - R_0$ is also called "excess risk". It measures the effect of an exposure in *absolute* terms. Its baseline is 0.

The **Relative Risk** $RR = \frac{R_1}{R_0}$ measures effect in relative terms. Its baseline is 1.

2-by-2 notation is common:

	D+	D−	Total
E+ (Group 1)	A_1	B_1	N_1
E− (Group 0)	A_0	B_0	N_2
	M_1	M_0	N

For person-time data, ignore cells B_1 and B_0 and let N_i represent group person-time. Thus,

$RD = R_1 - R_0 = (A_1 / N_1) - (A_0 / N_0)$, $RR = \frac{R_1}{R_0} = \frac{A_1 / N_1}{A_0 / N_0}$, and the **Odds Ratio**

$OR = \frac{A_1 B_0}{A_0 B_1}$. The OR is a good measure of relative effect in its own right and is equal to the RR when the outcome is rare (risks < 5%). The OR is the *only* legitimate measure of association in case-control studies.

The **Attributable Fraction in exposed cases** $AF_e = \frac{RR - 1}{RR}$ is the fraction of exposed cases that would be averted if they had not been exposed. Equivalent formulas are shown in the text, but the aforementioned formula is easiest and most versatile.

The **Attributable Fraction in the entire population** $AF_p = \frac{R - R_0}{R}$ reflects the potential reduction in morbidity associated with elimination of a harmful exposure. Equivalent formulas include $AF_p = \frac{(p_e)(RR-1)}{1+(p_e)(RR-1)}$ where p_e represents the proportion of the population that is exposed and *RR* represents the relative risk associated with exposure, and $AF_p = AF_e \times p_c$ where p_c represents the proportion of cases that are exposed to the risk factor. For protective factors, the equivalent measure is the **Preventive Fraction in the population** $PF_p = \frac{R_0 - R}{R_0}$ or, equivalently, $PF_p = \frac{p_c(1-RR)}{p_c(1-RR)+RR}$, representing the ppotential reduction in morbidity associated with universal introduction of the preventive exposure.

7: Rate Adjustment ("Standardization")

For uniformity of language, "rate" will be used to refer to any incidence or prevalence measure.

Direct Standardization

The directly adjusted rate (aR_{direct}), is a weighted average of strata-specific rates with weights coming from a reference population:

$$aR_{direct} = \sum w_i r_i$$

where $w_i = \frac{N_i}{\sum N_i}$ and

N_i represents the size of strata *i* of the reference population. (Capital letters denote values from the reference population)

r_i represents rate in strata *i* of the study population. (Lower case letters denote values from the study population.)

Indirect Standardization

The **Standardized Mortality Ratio (SMR)**

$$\text{SMR} = \frac{\text{Observed}}{\text{Expected}}$$

The "Observed" is merely the observed number of cases and the $\text{Expected} = \sum R_i n_i$ where R_i represents the rate in strata *i* of the reference population and n_i represents the number of people stata *i* of the study population. This formula can be understood in terms of the expected number of cases in strata *i*, which is merely $Expected_i = R_i n_i$.

The SMR is a fundamental population based relative risk estimate, with "1" representing a population with an observed rate that is as expected.

You can use the SMR to derive the indirectly adjusted rate $aR_{indirect} = (\text{crude rate}) \times SMR$.

Chapter 4: Screening for Disease

Reproducibility statistics

Rater B

Rater A	+	–	
+	a	b	p_1
–	c	d	q_1
	p_2	q_2	N

$$\kappa = \frac{2(ad - bc)}{p_1 q_2 + p_2 q_1}$$

Despite what the book says, you do not have to convert counts into percents to use this formula. κ quantifies the percent agreement that is above that due to chance. See Table 4.4 (p. 82) for interpretation guidelines.

Validity statistics

	Disease +	Disease –	Total
Test +	TP	FP	TP + FP
Test –	FN	TN	FN + TN
Total	TP + FN (those with disease)	FP + TN (those w/out disease)	*N*

SEN = (TP) / (those with disease) [note: TP = (SEN)(TP + FN)]
= (TP) / (TP + FN)

SPEC = (TN) / (those without disease) [note: TN = (SPEC)(FP + TN)]
= (TN) / (TN + FP)

PVP = (TP) / (those who test positive)
= (TP) / (TP + FP)

PVN = (TN) / (those who test negative)
= (TN) / (TN + FN)

True prevalence = (TP + FN) / N [True prevalence also known as *prior probability*]

See text for Bayesian equivalents for determining predictive value based on prior probabilities and test parameters.