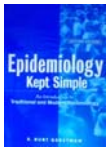
Epidemiology Kept Simple



Chapter 7

Rate Adjustment

Confounding

- ***Confounding*** ≡ a systematic error in inference due to extraneous factors
- **Potential confounder** ≡ an **extraneous factor** that can cause confounding
- Statistical techniques exist to adjust for confounding
- One such technique is **rate adjustment** (also called **rate standardization**)

Terminology

- For **uniformity of language**, let ***"rate"*** ≡ any incidence or prevalence, denoted *R* or *r*
- ***Crude rate*** (synonym: unadjusted rate) ≡ rate for entire population, denoted *cR*
- ***Strata-specific rate*** ≡ rate within a strata (subgroup), denoted r_i, for rate within stratum *i*
- ***Adjusted rate*** (synonym: *standardized rate*) ≡ mathematically compensated rate, denoted *aR*

Mortality *per 1000* by State, 1991

		Alaska		Florida	
i	Age	Deaths	Pop. (×1000)	Deaths	Pop. (×1000)
1	0–4	122	57	2,177	915
2	5–24	144	179	2,113	3,285
3	24–44	382	222	8,400	4,036
4	45–64	564	88	21,108	2,609
5	65–74	406	16	30,977	1,395
6	75+	582	7	71,483	1,038
	TOTAL	2,200	569	136,258	13,278

Crude rate, Alaska $cR_{Alask.} = \frac{2200}{569} = 3.9$

Crude rate, Florida $cR_{Florida} = \frac{136,258}{13,278} = 10.3$

Strata-Specific Rates by State

i	Age	Alaska	Florida
1	0–4	2.14	2.38
2	5–24	0.80	0.64
3	25–44	1.72	2.08
4	45–64	6.40	8.09
5	65–74	25.38	22.21
6	75+	83.14	68.87

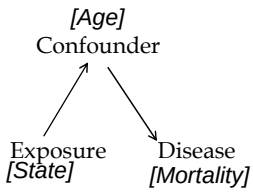
Age-specific rates similar to each other
Why did Florida have a much higher crude rate?

Age Distributions

Age	AK	%	FL	%
0-4	57	10%	915	7%
5-24	179	31%	3285	25%
25-44	222	39%	4036	30%
45-64	88	15%	2609	20%
65-74	16	3%	1395	11%
>75	7	1%	1038	8%
TOTAL	569	100%	13278	100%

Properties of Confounding

- Exposure associated with the confounder
- Confounder is a risk factor
- Confounder is not intermediary in the causal pathway



(c) B. Gerstman 2007 Chapter 7: Age Adjustment 7

What can we do about confounding?

- Like-to-like (strata-specific) comparisons (e.g., 80-year olds to 80-year olds)
- Mathematical adjustments:
  - Direct and indirect standardization (Ch 7)
  - Mantel-Haenszel techniques (Ch 14)
  - Logistic, Poisson, or Cox regression models (next edition)

(c) B. Gerstman 2007 Chapter 7: Age Adjustment 8

§7.2 Direct Adjustment

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

where

$$w_i = \frac{N_i}{N}$$

$N_i \equiv$ reference population size, strata i

$N \equiv$ reference population total size

$r_i \equiv$ rate, study population, strata i

aR_{direct} is a weighted average of strata-specific rates

(c) B. Gerstman 2007 Chapter 7: Age Adjustment 9

"Standard Million" 1991 Reference

Weight, strata *i* (w_i) = proportion in reference pop = N_i / N

i	Age	N_i	w_i
1	0–4	76,158	0.076158
2	5–24	286,501	0.286501
3	24–44	325,971	0.325971
4	45–64	185,402	0.185402
5	65–74	72,494	0.072494
6	75+	53,474	0.053474
	Σ→	*N* = 1,000,000	1.000000

(c) B. Gerstman 2007 Chapter 7: Age Adjustment 10

Alaska, Direct Adjustment

(Rates are per 1000)

i	Age	Rate r_i	Weights w_i	Product $w_i \cdot r_i$
1	0–4	2.14	0.076158	0.16297814
2	5–24	0.80	0.286501	0.22920080
3	24–44	1.72	0.325971	0.56067012
4	45–64	6.40	0.185402	1.18657280
5	65–74	25.38	0.072494	1.83989772
6	75+	83.14	0.053474	4.44582836
			$\Sigma w_i \cdot r$ =	8.42514792

$$aR_{Alask.} = \sum w_i r_i = 0.163 + 0.223 + \cdots + 4.45 \approx 8.43$$

Florida, Direct Adjustment

(Rates are per 1000)

i	Rate r_i	Weights w_i	Product $w_i \cdot r_i$
1	2.38	0.076158	0.18126
2	0.64	0.286501	0.18336
3	2.08	0.325971	0.67802
4	8.09	0.185402	1.49990
5	22.21	0.072494	1.61009
6	68.87	0.053474	3.68275
		$\Sigma w_i \cdot r$ =	7.83538

$$aR_{Florida} = 0.181 + 0.183 + \cdots + 3.683 = 7.84$$

Conclusions

- cR_{FL} (10.3) > cR_{AK} (3.9)
- aR_{FL} (7.84) < aR_{AK} (8.47)
- Age confounded the crude comparison
  - State (E) associated with age (C)
  - Age (C) is independent risk factor for death (D)
  - Age (C) is not an intermediary in the causal pathway between E and D

§7.3 Indirect Adjustment

This method is based on a statistic called the Standardized Mortality Ratio (*SMR*):

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

where

$$\text{Observed} = \text{observed no. of cases}$$

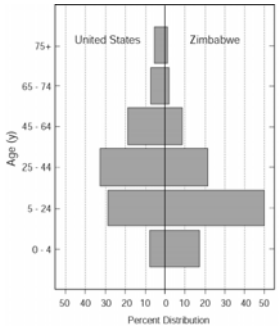
$$\text{Expected} = \text{expected no. of cases}$$

Example: Zimbabwe & US

Crude death rate in US (1991)
= 880 per 100,000

Crude death rate in Zimbabwe
= 886 per 100,000

But, US population much older



Zimbabwe

Observed number of deaths = 98,808
Expected number calculated as follows:

i	Age	US R_i	Zimbabwe n_i	$\mu_i = R_i \times n_i$
1	0–4	00229	1,899,204	4,349
2	5–24	.00062	5,537,992	3434
3	24–44	.00180	2,386,079	4,295
4	45–64	.00789	974,235	7,687
5	65–74	.02618	216,387	5,665
6	75+	.08046	136,109	10,951

$$\text{Expected} = \sum \mu_i = 4349 + 3434 + \ldots + 10{,}951 = 36{,}381$$

Zimbabwe SMR

$$SMR = \frac{\text{Observed}}{\text{Expected}} = \frac{98{,}808}{36{,}381} = 2.72$$

After adjusting for age, Zimbabwe's mortality rate is 2.72 times that of the U.S.

Indirect Adjustment (optional)

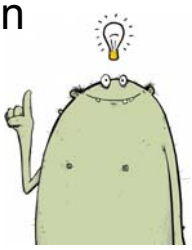
$$aR_{indirect} = (cR)(SMR)$$

- Zimbabwe crude rate = 886 (per 100,000)
- aR_{indirect} = cR × SMR = (860)(2.72) = 2340
- Not necessary: interpret the SMR directly (instead)

7.4: Adjustment for Multiple Factors

- Same methods, just stratify by all factors
- For example, to control for age & gender stratify you could stratify as follows
  - 0-4 females
  - 0-4 males
  - 5-9 females
  - Yada yada yada
- … then apply standardization method

Conclusion



- Crude rates cannot be used when populations are "not comparable"
- Simplest method for fair comparison is to restrict to strata-specific comparisons
- Mathematical methods of adjustment exist, including direct and indirect standardization