JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

The Direct and Indirect Methods of Adjustment

Marie Diener-West, PhD
Sukon Kanchanaraksa, PhD
Johns Hopkins University

JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Section A: Introduction

Marie Diener-West, PhD
Johns Hopkins University

Comparing Rates Among Different Populations

- **Stratum-specific rates** (e.g., age-specific rates)
- **Crude rates**
- **Stratum-adjusted rates** (e.g. age-adjusted rates)

Comparison of Stratum-Specific Rates

- If a population can be **stratified** (divided into groups), appropriate comparisons may be made of **stratum-specific rates** such as:
  - Age-specific rates
  - Cause-specific rates
  - Age-cause-specific rates
  - Age-gender-race-specific rates

Comparison of Crude Rates

- A **crude rate** (overall rate) is a weighted average of stratum-specific rates (the weights are the population totals of the strata)
- The difference between crude rates of two populations involves differences in both the stratum-specific rates and population composition (distribution of characteristics)
- Comparison of crude rates is often confounded by these differences and not appropriate

Notation

- Let
  - j = stratum
  - x_j = number of events in stratum j of the observed population
  - N_j = number of persons in stratum j of the observed population
  - p_j = event rate in stratum j of the observed population
  - C = crude rate of the observed population

$$C = \frac{\Sigma x_j}{\Sigma N_j} = \frac{\text{total deaths}}{\text{total population}}$$

- Stratum-specific rate $= \frac{X_j}{N_j} = \frac{\text{deaths in stratum j}}{\text{population in stratum j}}$

Example: Mortality by Age in Population 1

Population 1			
Stratum i (Age group)	Age (years)	Total population (N_i)	Deaths (X_i)
1	0–4	100	5
2	5–14	90	10
3	15–19	110	15
Total		300	30

- **Crude death rate** = 30 deaths/300 population = 0.10 = 10 deaths per 100 population

Example: Mortality by Age in Population 2

Population 2			
Stratum i (Age group)	Age (years)	Total population (N_i)	Deaths (X_i)
1	0–4	165	10
2	5–14	75	10
3	15–19	60	10
Total		300	30

- **Crude death rate** = 30 deaths/300 population = 0.10 = 10 deaths per 100 population

Example: Comparison of Crude Death Rates

- The **crude death rate (CDR)** for each population is 10 deaths per 100 population
- Is the **risk of dying** the same in the two populations?

Example: Mortality by Age in Population 1

Population 1				
Stratum i (Age group)	Age (years)	Total Population (N_{1i})	Deaths (X_{1i})	Death Rate per 100 (X_{1j}/N_{1j})*100
1	0–4	100	5	5
2	5–14	90	10	11
3	15–19	110	15	14
Total		300	30	

Example: Mortality by Age in Population 2

Population 2				
Stratum i (Age group)	Age (years)	Total Population (N_{2i})	Deaths (X_{2i})	Death Rate per 100 (X_{2j}/N_{2j})*100
1	0–4	165	10	6
2	5–14	75	10	13
3	15–19	60	10	17
Total		300	30	

Example: Comparison of Age-Specific Death Rates

- Age-specific death rates (per 100 population)

Stratum i	Age	Population 1	Population 2
1	0–4	5	6
2	5–14	11	13
3	15–19	14	17

Conclusion: Age-Specific vs. Crude Death Rates?

- Inspection of the age-specific rates reveals a **higher death rate in each age group** for Population 2 as compared to Population 1
- **Why** are the **crude rates the same** in the two populations?
  - Hint: look at the age composition (population distribution by age)!

Population 1 Distribution

Population 1				
Stratum i	Age	Total	%	Death rate per 100
1	0–4	100	33	5
2	5–14	90	30	11
3	15–19	110	37	14

Population 2 Distribution

Population 2				
Stratum i	Age	Total	%	Death rate per 100
1	0–4	165	55	6
2	5–14	75	25	13
3	15–19	60	20	17

Crude Death Rates for Populations 1 vs. 2

- The **Crude Death Rate (CDR)** is a **weighted average** of the age-specific death rates.

- For Population 1:

$$CDR_1 = \frac{\Sigma p_{1j} N_{1j}}{N_{1j}} = \frac{(0.05 \times 100) + (0.11 \times 90) + (0.14 \times 110)}{300} = .10$$

- For Population 2:

$$CDR_2 = \frac{\Sigma p_{2j} N_{2j}}{N_{2j}} = \frac{(0.06 \times 165) + (0.13 \times 75) + (0.17 \times 60)}{300} = .10$$

What Is the Appropriate Comparison?

- The comparison of crude death rates is confounded by the **differences in population composition** (age distribution) between the two populations
- Population 2 has a younger age distribution
- A comparison of age-specific death rates between the two populations reflects the risk of dying in **each age group**
- An **adjustment procedure** is needed to make an appropriate comparison of the **overall** risk of dying between the two populations

Review

- Why might crude death rates be misleading?
- What is an alternative to comparing crude death rates?

Next Steps

- The two most common adjustment procedures for rates:
  - **Direct method** of adjustment
  - **Indirect method** of adjustment
- Commonly used in vital statistics and epidemiology.

JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

Section B: Direct Method of Adjustment

Marie Diener-West, PhD
Johns Hopkins University

Adjustment Procedures

- **Adjustment procedures** are any of a variety of procedures performed during data analysis to attempt to remove the effect of one or more extraneous sources of variation that could affect (or are believed to affect) a particular result (Meinert, 1996)

Goals of Adjustment Procedures

- Appropriately combine data
- Make appropriate comparisons among groups
- Reach appropriate conclusions and inferences

Types of Adjustment Procedures

- Stratified or subgroup analyses
- Direct or indirect standardization of rates
- Life tables
- Multivariable statistical analyses

Adjusted or Standardized Rates

- Calculation of **adjusted (standardized) rates** allows comparison of summary event rates between populations when there are differences in characteristics between the populations that may influence the event of interest
  - For example, age, race, gender, disease status

Methods of Adjustment of Rates

- **Direct method**
  - Apply stratum-specific rates observed in the populations of interest to a **reference** or **standard population** in order to obtain the number of deaths expected in the reference population
  - Calculate an adjusted rate based on expected number of deaths in the reference population
- **Indirect method**
  - Apply stratum-specific **reference rates** to the populations of interest to obtain the number of expected deaths in each of those populations
  - Compare the observed number of deaths to the expected number of deaths for each population of interest

Notation for Populations of Interest

N_{1j} =	Number of individuals in stratum j of Population 1
N_{2j} =	Number of individuals in stratum j of Population 2
X_{1j} =	Number of individuals in stratum j of Population 1 who *have* the event
X_{2j}	Number of individuals in stratum j of Population 2 who *have* the event

$P_{1j} = X_{1j} / N_{1j} =$	Rate in stratum j of Population 1
$P_{2j} = X_{2j} / N_{2j} =$	Rate in stratum j of Population 2

Notation for Reference Population

N_j =	Number of individuals in stratum j of a reference population
P_j =	Rate for stratum j of the reference populations
C =	Crude rate of the reference population

Adjusted Rates by the Direct Method

- Use the **event rates (experience) of the population of interest** to calculate the number of deaths expected in the reference population
- The **adjusted rate** in the population of interest is the expected number of deaths divided by the total reference population

Adjusted Rates by the Direct Method

- Adjusted rate for Population 1

$$= \frac{\sum p_{1j} \cdot N_j}{\sum N_j}$$

- Adjusted rate for Population 2

$$= \frac{\sum p_{2j} \cdot N_j}{\sum N_j}$$

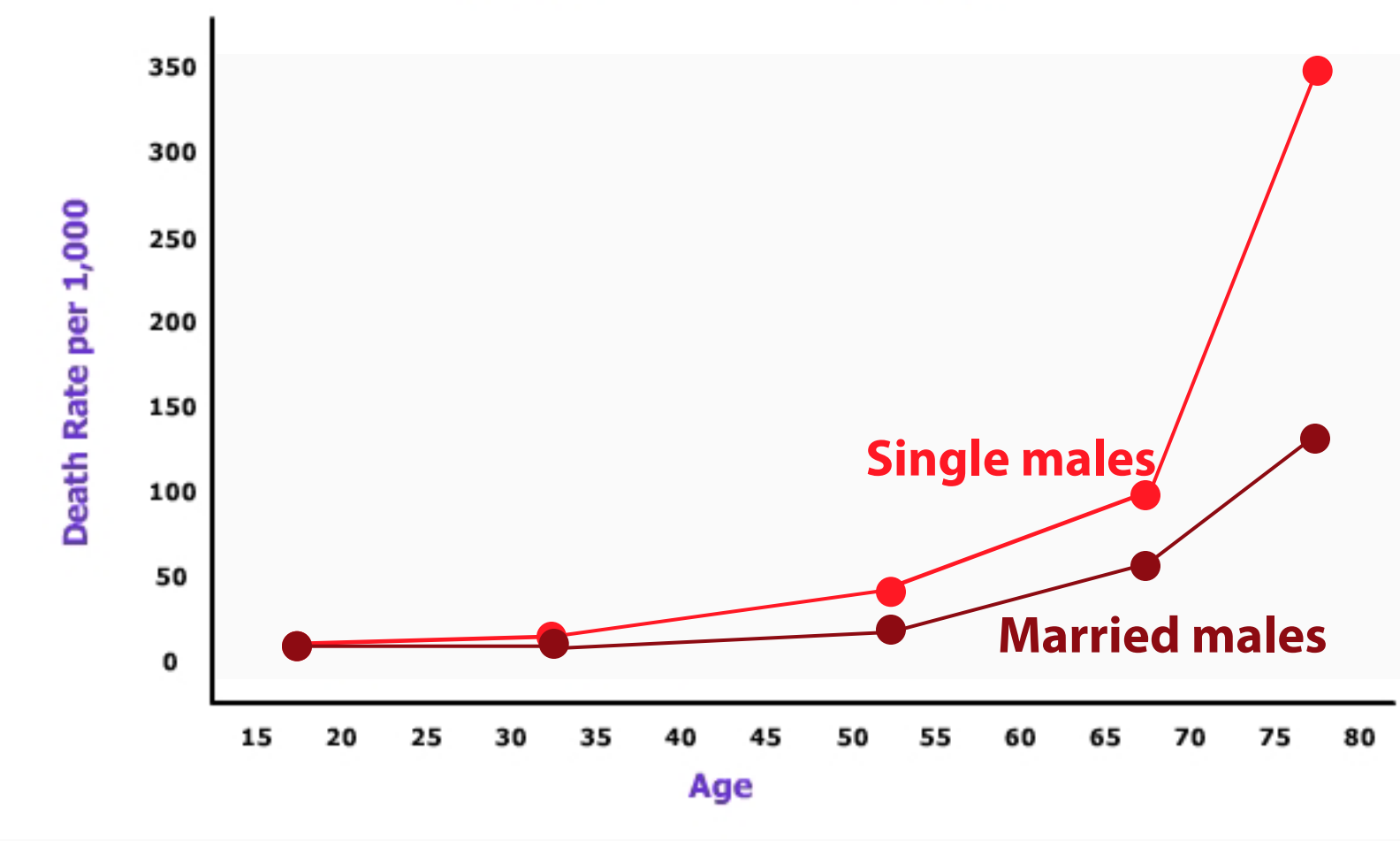
Example: U.S. Mortality—Single vs. Married Males

		Single Males			Married Males		
		Column 1	**Column 2**	**Column 3**	**Column 4**	**Column 5**	**Column 6**
J	Age	**N_{1j} Population (thousands)**	**X_{1j} Number of Deaths**	**P_{1j} Death Rate Per 1,000**	**N_{2j} Population (thousands)**	**X_{2j} Number of Deaths**	**P_{2j} Death Rate Per 1,000**
1	15–24	17,724	39,745	2.24	3,427	7,327	2.14
2	25–44	5,390	26,372	4.89	23,083	60,337	2.61
3	45–64	1,210	37,125	30.68	18,088	286,535	15.84
4	65–74	364	33,679	92.52	5,500	275,818	50.15
5	75+	199	64,386	323.55	2,331	275,135	118.03
Total		**24,887**	**201,307**	**8.09**	**52,429**	**905,152**	**17.26**

Crude Mortality Rates: Single vs. Married Males

- **Crude death rates** per 1,000 population are:
  - Single: 8.09
  - Married: 17.26
- A comparison of crude rates gives the impression that the risk of dying is twice as high in married males as single males

U.S. Males, Death Rates by Marital Status and Age



Summary: U.S. Mortality, Single vs. Married Males

- The crude death rate is higher for married males than single males
- However, age-specific death rates per 1,000 are lower for married males in each age stratum
- Why is this?
- The two populations have **very different age distributions:**
  - Single: 90% are aged < 45
  - Married: 51% are aged < 45

Age-Specific Death Rates: Single vs. Married Males

		Single Males			Married Males		
		Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
J	Age	N_{1j} Population (thousands)	X_{1j} Number of Deaths	P_{1j} Death Rate Per 1,000	N_{2j} Population (thousands)	X_{2j} Number of Deaths	P_{2j} Death Rate Per 1,000
1	15–24	17,724	39,745	2.24	3,427	7,327	2.14
2	25–44	5,390	26,372	4.89	23,083	60,337	2.61
3	45–64	1,210	37,125	30.68	18,088	286,535	15.84
4	65–74	364	33,679	92.52	5,500	275,818	50.15
5	75+	199	64,386	323.55	2,331	275,135	118.03
Total		**24,887**	**201,307**	**8.09**	**52,429**	**905,152**	**17.26**

Calculating Expected Deaths: U.S. Mortality, Males

		Column 3	Column 6	**Column 7**	Column 3 x Column 7	Column 6 x Column 7
J	Age	P_{1j} Single Death Rate per 1000	P_{2j} Married Death Rate per 1000	**N_j Reference Population (Thousands$_j$)**	Expected Deaths—Single	Expected Deaths—Married
1	15-24	2.24	2.14	**21,151**	47,430	45,221
2	25-44	4.89	2.61	**28,473**	139,312	74,426
3	45-64	30.68	15.84	**19,298**	592,097	305,703
4	65-74	92.52	50.15	**5,864**	542,565	294,072
5	75+	323.55	118.03	**2,530**	818,576	298,624
Total				**77,316**	2,139,980	1,018,046

Note: rates used to calculate expected deaths were carried to more decimal places
Reference population: total of single and married populations

Calculating Expected Deaths: U.S. Mortality, Males

		Column 3	Column 6	Column 7	**Column 3 x Column 7**	Column 6 x Column 7
J	Age	P_{1j} Single Death Rate per 1000	P_{2j} Married Death Rate per 1000	N_j Reference Population (Thousands)	**Expected Deaths—Single**	Expected Deaths—Married
1	15-24	2.24	2.14	21,151	**47,430**	45,221
2	25-44	4.89	2.61	28,473	**139,312**	74,426
3	45-64	30.68	15.84	19,298	**592,097**	305,703
4	65-74	92.52	50.15	5,864	**542,565**	294,072
5	75+	323.55	118.03	2,530	**818,576**	298,624
Total				77,316	**2,139,980**	1,018,046

Note: rates used to calculate expected deaths were carried to more decimal places
Reference population: total of single and married populations

Calculating Expected Deaths: U.S. Mortality, Males

		Column 3	Column 6	Column 7	Column 3 x Column 7	**Column 6 x Column 7**
J	Age	P_{1j} Single Death Rate per 1000	P_{2j} Married Death Rate per 1000	N_j Reference Population (Thousands)	Expected Deaths—Single	**Expected Deaths—Married**
1	15-24	2.24	2.14	21,151	47,430	**45,221**
2	25-44	4.89	2.61	28,473	139,312	**74,426**
3	45-64	30.68	15.84	19,298	592,097	**305,703**
4	65-74	92.52	50.15	5,864	542,565	**294,072**
5	75+	323.55	118.03	2,530	818,576	**298,624**
Total				77,316	2,139,980	**1,018,046**

Note: rates used to calculate expected deaths were carried to more decimal places
Reference population: total of single and married populations

Directly Age-Adjusted Death Rate: Single Males

- **Age-adjusted rate per 1,000 for single males**

$$= \frac{\text{expected deaths in reference population using single rates}}{\text{total reference population}}$$

$$= \frac{2{,}139{,}980}{77{,}316}$$

- DAR_s = 27.68 deaths per 1,000 population

Directly Age-Adjusted Death Rate: Married Males

- **Age-adjusted rate per 1,000 for married males**

$$= \frac{\text{expected deaths in reference population using married rates}}{\text{total reference population}}$$

$$= \frac{1{,}018{,}046}{77{,}316}$$

- DAR_m = 13.17 deaths per 1,000 population

Comparing Adjusted Rates: Married vs. Single Males

- The **directly age-adjusted death rates** per 1,000 are:
  - Single: $DAR_s = 27.68$
  - Married: $DAR_m = 13.17$
- A comparison of the directly age-adjusted rates reveals that the risk of dying is **twice as high for single versus married males** after adjusting for the differences in population age distribution between the two groups

Quick Check

- Why do the crude death rates suggest higher risk of dying in married males than in single males?

Notation

Let j = stratum

x_j = number of events in stratum j of the observed population

N_j = number of persons in stratum j of the observed population

p_j = event rate in stratum j of the observed population

$$DAR_i = \frac{\sum p_{1j} \cdot N_j}{\sum N_j}$$

$$DAR_i = \frac{\sum \text{expected deaths in reference population (using Pop}^{\underline{n}}\text{ rates)}}{\text{total reference population}}$$

Another Example of Age-Adjustment

Comparison of Crude Death Rates in a Population at Two Different Time Periods					
Early Period			Later Period		
Population	No. of Deaths	Rate per 100,000	Population	No. of Deaths	Rate per 100,000
900,000	862	96	900,000	1,130	126

An Example of Age-Adjustment

Comparison of Age-Specific Death Rates in the Two Time Periods						
Age Group	Early Period			Later Period		
	Population	No. of Deaths	Death Rate per 100,000	Population	No. of Deaths	Death Rate per 100,000
30–49	500,000	60	12	300,000	30	10
50–69	300,000	396	132	400,000	400	100
70+	100,000	406	406	200,000	700	350
Total	900,000	862	96	900,000	1,130	126

An Example of Age-Adjustment

Carrying out an Age Adjustment Using the Total of the Two Populations as the Standard					
Age Group	Standard Population	Early Period		Later Period	
		Death Rate per 100,000	Expected No. of Deaths	Death Rate per 100,000	Expected No. of Deaths
30–49	800,000	12	96	10	80
50–69	700,000	132	924	100	700
70+	300,000	406	1,218	350	1,050
Total	**1,800,000**		**2,238**		**1,830**
Age-Adjusted Rate			$\frac{2,238}{1,800,000}$		$\frac{1,830}{1,800,000}$

Summary of the Direct Method of Adjustment

- **The age-adjusted rate** is an index measure, the magnitude of which has no intrinsic value
- The actual rates are only meaningful when directly compared to each other
- The adjusted rates are useful for comparison purposes only
- The choice of the reference population is important
  - It should not be abnormal or unnatural
- Adjustment (standardization) is **not** a substitute for the examination of age-specific rates in the populations of interest

Review Questions

- What is an age-adjusted rate?
- Can directly-adjusted rates be compared?
- Using the direct method of adjustment, what data are needed from the reference (standard) population?

JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Section C: Indirect Adjustment

Sukon Kanchanaraksa, PhD
Johns Hopkins University

Direct versus Indirect Method of Adjustment

- Direct method
  - Assume that the observed population had the same distribution of characteristics as the reference population
  - Apply the experience of the observed population to the reference population
- Indirect method
  - Assume that the observed population had the experience of the reference population
  - Apply the experience of the reference population to the observed population

Experience = morbidity rate or mortality rate

Characteristics = age, gender, …

Indirect Method of Adjustment

- Apply the experience (e.g., mortality rate) of the reference (standard) population to the population of interest (observed population) by strata (age group)
- Sum the values to obtain the expected number of deaths
- Divide the observed number of deaths by the expected number of deaths in the population of interest to get a value called **SMR (Standardized Mortality Ratio)**
- Multiply SMR by the **crude rate (C)** of the reference population to get the **Indirect Adjusted Rate (IAR)**

Notation

Let j = strata (e.g., age group)

x_j = number of events in stratum j of the observed population (e.g., observed number of deaths)

N_j = number of persons in stratum j of the observed population

p_j = rate in stratum j of the reference population (e.g., mortality rate)

C = crude rate of the reference population

$$SMR = \frac{\Sigma X_j}{\Sigma(p_j N_j)}$$

$$IAR = SMR \times C$$

Thus: $$SMR = \frac{IAR}{C}$$

Example of an Indirect Adjustment

- In a population of 534,533 White male miners, 436 died from tuberculosis (TBC) in 1950
- Is this mortality experience from TBC **greater than, less than, or about the same as** that which you would expect in White males of the same ages in the general population?
- And, what is the IAR of the White male miners compared to the general population of White males?

Computation of an SMR for Tuberculosis

Computation of an SMR for Tuberculosis, All Forms (TBC), for White Male Miners Ages 20–59 yrs, U.S., 1950				
(j) Age (Yrs)	(N_j) Estimated Population of White Miners (1)	(p_j) Death Rate Per 100,000 for TBC in White Males (General Population) (2)	(p_j .N_j) Expected Deaths from TBC in White Miners (If Same Risk as General Population (3)=(1)X(2)	(X_j) Observed Deaths from TBC in White Miners (4)
20–24	74,598	12.26	9.14	10
25–29	85,077	16.12	13.71	20
30–34	80,845	21.54	17.41	22
35–44	148,870	33.96	50.55	98
45–54	102,649	56.82	58.32	174
55–59	42,494	75.23	31.96	112
Total			181.09	436

Adapted from: Vital Statistics - Special Reports, DHEW, 53(5), 1963

Computation for SMR and IAR

- $\text{SMR} = \dfrac{\text{observed deaths}}{\text{expected deaths}}$

  $= \dfrac{436}{181.09} \times 100$

  $= 241\% \quad \text{or} \quad 2.41$

- $\text{IAR} = \text{SMR} \times C$

  $= 2.41 \times 10.9 \text{ per } 1000$

  $= 26.3 \text{ per } 1000$

Interpretation of SMR

- SMR = 1
  - Risk is the same in both the observed population and the reference population
- SMR < 1
  - Risk is lower in the observed population compared to the reference population
- SMR > 1
  - Risk is higher in the observed population compared to the reference population

SMR = 2.41 $\Rightarrow$ White miners had 2.41 times the risk of mortality of the U.S. White male population

The indirect-adjusted mortality rate for White miners was 26.3 per 1000

Comparison of SMRs

- Since the number of deaths in a population depends on age distribution, the number of observed deaths and the calculation of the expected deaths must depend on the age distribution of the population of interest
  - Consequently, SMR must also depend on the age distribution of population of interest
- Therefore, when using the same reference population in the calculation of SMR, the SMR from one population cannot be compared to the SMR from another population **unless** the two populations are similar in age distribution

Quick Check

- If the SMR from a textile worker industry were 4.0 or 400% and the SMR from miners were 2.4 …
  - Could we conclude that the risk of death of the textile workers was 4/2.4=1.7 times higher than the miners?

Example of SMR by Occupation

SMR by Occupation Men Aged 20–64 in the United States, 1950		
	SMR	
Occupation Level	Disease of Respiratory System	Asthma
All occupations	100	100
Professional workers	72	71
Technical, administrative and management workers	52	79
Clinical sales and skilled workers	87	104
Semiskilled workers	149	99
Laborers	157	145
Agricultural workers	75	95

Source: Higgins, 1974

Merits of Indirect Adjustment

- No need to know the age-specific (mortality) rates of the population of interest
  - These rates may be difficult to obtain (in a developing country or an industry)
  - The direct method of adjustment cannot be conducted without these rates
- The rates of the standard population are often based on large population, while the rates calculated from the observed population may be based on small number and be unstable

Review Questions

- What is SMR?
- Can SMR of one population be compared to SMR of another population?
- In the indirect method, what data from the reference (standard) population is used?