Foreword

This report has been published jointly by the Public Health Commission (PHC) and the Ministry of Health and it aims to provide health practitioners, data analysts, researchers and policy makers with some basic and essential tools for analysing health statistical information.

When analysing health issues much of the statistical information we first encounter is presented to us in the form of numbers and sometimes crude rates. The trends and patterns we need to examine in these statistics often interact with social and demographic factors such as age, gender, ethnicity and social class. This report provides a method which assists to account for the effect of these factors in analyses of health issues.

The report was prepared by Dr Barry Borman, a senior epidemiologist in the PHC. We acknowledge also the work of Dr Neil Pearce of the Wellington School of Medicine who generously reviewed the draft report.

Standardising Rates of Disease is a third report in an Analysis & Monitoring series to be published by the PHC. The report is published jointly by the PHC and the Ministry of Health and revises a previous edition published by the Department of Health in 1992. It is based on notes used for teaching standardisation in the Department of Public Health at the Wellington School of Medicine, the Department of Nursing Studies at Victoria University, and the Ministry of Health, Wellington. A computer disk containing spreadsheets of all the examples and exercises in Microsoft EXCEL format is available on request from the author. We would be pleased to receive comments and suggestions about this and future reports, and readers are invited to write to the Public Health Commission, PO Box 1795, Wellington.

Paul Brown
Group Director
Analysis and Monitoring
Public Health Commission

Judith Johnston
Deputy Director-General
New Zealand Health Information Service
Ministry of Health

Contents

Foreword	1
Introduction	3
Notation	4
Example Data	5
Crude Rate	6
Age-specific Rate	8
Standardisation	9
Direct Standardisation	11
Issues in the Use of Direct Standardisation	12
Indirect Standardisation	16
Issues in the Use of Indirect Standardisation	18
Standard Population and Rates	20
Some Standard Populations Often Used in Standardisation	21
The Effect of Using Different Standard Populations	22
Issues in the Use of Standardisation	23
Reminder	24
References	24
Exercises	25
Answers to the Exercises	28

Introduction

A frequent problem in analysing health data is to make valid comparisons of the rate at which some health event (eg, death from various causes) occurs in different populations or over time. The populations are rarely similar with respect to demographic factors associated with the health event being studied - such as age, ethnicity and social class.

This report describes a method for “standardising” rates of disease which “adjusts” or “controls” for such demographic differences between populations. It is assumed that standardisation is used to adjust for the possible effects of age on the rates of disease. The method may also be used to control for the effects of various other demographic factors such as social class, ethnicity or parity.

Notation

	Age group (years)			
	<25	25-64	65+	Total
		Study Group		
number of events per annum (deaths or cases) *numerator*	n_1	n_2	n_3	Σn_i
population-at-risk or person-years-at-risk *denominator*	p_1	p_2	p_3	Σp_i
rate	$asr_1 = \frac{n_1}{p_1} * C$	$asr_2 = \frac{n_2}{p_2} * C$	$asr_3 = \frac{n_3}{p_3} * C$	$cr = \frac{\Sigma n_i}{\Sigma p_i} * C$
		Standard Population		
number of events per annum (deaths or cases) *numerator*	N_1	N_2	N_3	ΣN_i
population-at-risk or person-years-at-risk *denominator*	P_1	P_2	P_3	ΣP_i
rate	$ASR_1 = \frac{N_1}{P_1} * C$	$ASR_2 = \frac{N_2}{P_2} * C$	$ASR_3 = \frac{N_3}{P_3} * C$	$CR = \frac{\Sigma N_i}{\Sigma P_i} * C$

Σ = total
asr/ASR = age-specific rate
cr/CR = crude rate
C = constant or unit of population

Example Data

	Age group (years)			
	<25	25-64	65+	Total
		Study Group A		
deaths (*numerator*)	5	9	20	34
population (*denominator*)	1,000	900	500	2,400
rate (per 1,000 population)	5.0[1]	10.0[1]	40.0[1]	14.2[2]
		Study Group B		
deaths (*numerator*)	5	10	40	55
population (*denominator*)	1,000	1,000	1,000	3,000
rate (per 1,000 population)	5.0[1]	10.0[1]	40.0[1]	18.3[2]
		Standard Population		
deaths (*numerator*)	50	90	96	236
population (*denominator*)	10,000	9,000	8,000	27,000
rate (per 1,000 population)	5.0[1]	10.0[1]	12.0[1]	8.7[2]

[1] Age-specific rate per 1,000 population

[2] Crude rate per 1,000 population

Crude Rate

A crude, or "unadjusted", rate (CR) is a measure of the actual mortality or morbidity experience of the population under study (ie, the "population-at-risk"). It is calculated by dividing the total number of deaths or cases (ie, *the numerator*) in the study population, by the total population-at-risk or person-years-at-risk (ie, *the denominator*), expressed as a unit of the population (eg, per 1,000 population).

The formula for a crude rate

(using the notation on page 4)

$$CR \text{ (per 1,000 population)} = \frac{\Sigma n_i}{\Sigma p_i} * 1{,}000 \text{ (the constant or unit of population)}$$

Where:

n_i = the total number of deaths or cases (numerator) in age group i per annum
p_i = the total population (denominator) in age group i

Example

Using the notation and data on pages 4 and 5, the crude rate per 1,000 people in study group A (34 deaths in a population of 2,400 people) is:

$$CR \text{ per } 1{,}000 = \frac{\Sigma n_i}{\Sigma p_i} = \frac{n_1 + n_2 + n_3}{p_1 + p_2 + p_3} * 1{,}000$$

$$= \frac{5 + 9 + 20}{1{,}000 + 900 + 500} = \frac{34}{2{,}400} * 1{,}000$$

$$= 14.2 \text{ per } 1{,}000 \text{ population}$$

The crude rate in study group B (55 deaths in a population of 3,000 people) is:

$$= \frac{5 + 10 + 40}{1{,}000 + 1{,}000 + 1{,}000} = \frac{55}{3{,}000} * 1{,}000$$

$$= 18.3 \text{ per } 1{,}000 \text{ population}$$

Note:

a) The crude rate is a summary of the total actual mortality or morbidity experience in a population.

b) The crude rate is influenced by the distribution of the population-at-risk (or denominator) in the various age groups of the study group. In the example above, 42 percent of the population in study group A was under the age of 25 years, compared to 33 percent in study group B. Conversely, 33 percent of the people in study group B were over 65 years of age compared to 21 percent in study group A. As the rate of many diseases increases with age the crude rate for study group B is the highest because it has a higher proportion of older people than study group A.

The effect of differences in population structures on crude rates

Table 1: Rates of Down Syndrome Reported in Various Studies

Area	Crude rate[1]	Crude rank	Age-adjusted rate[2]	Age-adjusted rank
Jerusalem (Israel)	2.4	1	1.4	1
Milan (Italy)	1.6	2	0.9	8
London (England)	1.5	3	0.9	9
Victoria (Australia)	1.4	4	0.9	10
Massachusetts (USA)	1.4	5	1.0	7
Sweden	1.3	6	1.1	2
British Columbia (1952-59)	1.3	7	0.9	11
Copenhagen (Denmark)	1.2	8	1.1	3
Australia	1.1	9	1.0	4
British Columbia (1972-75)	1.0	10	1.0	5
Atlanta (USA)	1.0	11	1.0	6

Source: Adams, MM, Erickson JD, Layde PM, et al. Down syndrome: recent trends in the United States, JAMA, 1981; 246: 758-60.

1 Rate per 1,000 livebirths

2 Rate per 1,000 livebirths adjusted for differences in the age distribution of the maternal populations

Note:

a) A strong relationship has been shown between the occurrence of Down syndrome and maternal age with older women most at-risk to giving birth to an infant with Down syndrome. Therefore, the crude rate of Down syndrome is likely to be highest in populations which have a high proportion of women giving birth over the age of 35 years.

b) In *Table 1* there is a marked change in the crude rate and the rate adjusted for differences in the maternal age distribution of the various population groups (eg, Jerusalem, Milan and London). There was no change in the rate for Atlanta because the maternal population of Atlanta was used as the standard population (the so-called direct method was used for standardising these rates).

Age-specific Rate

An age-specific rate (ASR) is the rate at which a particular health event (eg, death or disease incidence) occurs in each age group of a population, expressed as some unit of the population-at-risk or person-years-at-risk. (Remember the technique described in this handbook can also be used to calculate other factor-specific rates such as for ethnic group and social class).

An age-specific rate is simply the crude rate for the specific age group. For example, to calculate the age-specific rate of a disease for people aged 45 to 49 the total number of cases in the age group is divided by the population in that age group and multiplied by a constant (a unit of population, such as 1,000 or 100,000).

The formula for an age-specific rate

(using the notation on page 4)

$$\text{ASR per 1,000 population} = \frac{n_i}{p_i} * 1{,}000 \text{ (constant)}$$

Where:

n_i = the number of deaths or cases (numerator) in age group i per annum

p_i = the total population (denominator) in age group i per annum

Example

Using the data on page 5 for study group A, the age-specific rate (ASR) per 1,000 for the 25 to 64 age group is:

$$\text{ASR per 1,000 population} = \frac{n_2}{p_2} * 1{,}000$$

$$= \frac{9}{900} * 1{,}000$$

$$= 10.0 \text{ per 1,000 population}$$

The age-specific rate for the 65+ age group in study group B is:

$$\text{ASR (per 1,000)} = \frac{n_3}{p_3} * 1{,}000$$

$$= \frac{40}{1{,}000} * 1{,}000$$

$$= 40.0 \text{ per 1,000 population}$$

Note:

If there are small numbers of cases or deaths (numerator) or a small population (denominator) in some age groups, any age-specific rates that are calculated may be too imprecise and unreliable for use in detailed comparisons. For example, two deaths in a population of 25 will produce an age-specific rate of 80.0 per 1,000. One extra death would increase the rate to 120.0 per 1,000. In a population of 5,000 people there would have to be 400 deaths to produce a rate of 80.0, but 200 extra deaths would be needed to increase the rates to 120.0 per 1,000.

Standardisation

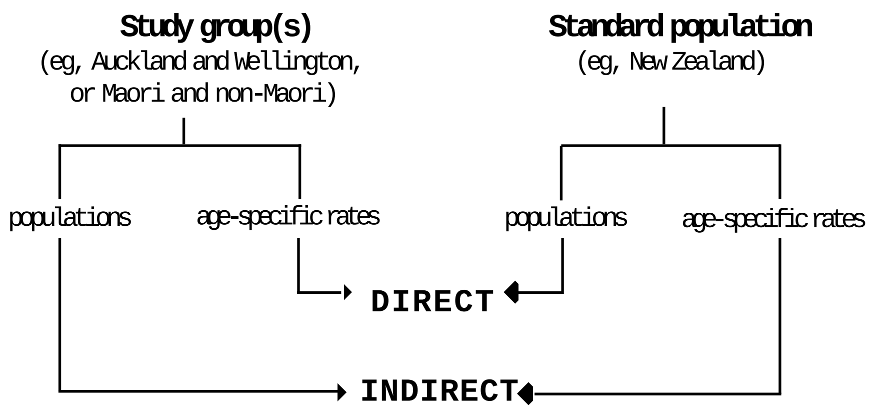
Comparing crude rates can lead to erroneous conclusions being drawn about the impact of a health event in a population. For example, *Table 1* (see page 7) has shown that differences between the crude rate of Down syndrome in study populations may be due to differences in the age distribution of mothers.

An age-specific rate is a measure of the disease or death in an age group, but it can be unreliable or imprecise if the numerator or denominator is small. Comparing the age-specific rates of two (or more) study populations can be difficult if there are a large number of age groups involved. For example, comparing five-year age-specific rates for Maori, Pacific Islands people, and "Others" involves making 54 comparisons (ie, 18 age groups and three ethnic groups).

It is, therefore, more convenient to compare the mortality or morbidity in study groups using a summary index which takes into account differences in the age distribution of the compared groups.

In analysing health data two methods of standardisation (direct and indirect) are commonly distinguished, but only one process is involved – age-specific rates are weighted by a population in specific age groups. The difference between direct and indirect standardisation is the source of the population (or weights) and the age-specific rates (see *Figure 1*).

Figure 1: Comparison of Direct and Indirect Standardisation



Assumptions and methods

Direct standardisation

- Assumes that the distribution of the population in the various age groups of the study group(s) is the same as that in a standard population.

- The age-specific rates of the study group (eg, Auckland and Wellington, Maori and non-Maori) are weighted (ie, multiplied) by, the population in each age group of the standard population (eg, New Zealand).

Indirect standardisation

- Assumes that the age-specific rates in the study group(s) are the same as in a standard population.

- The age-specific rates of the standard population (eg, New Zealand) are weighted (ie, multiplied) by, the population in each age group of the study group (eg, Auckland and Wellington, Maori and non-Maori) to estimate the "expected number" of events.

To use direct standardisation requires data on the number of deaths or cases and the number of people or person-years in each age group of the study population. Indirect standardisation is the preferred method when these data are not available or the population in the age groups is so small that there will be large fluctuation in the age-specific rates with the addition or absence of a few deaths or cases.

Direct Standardisation

The basis for using direct age-standardisation is to find out:

How many deaths or cases would have occurred in the study group if it had an identical age structure to some standard population, but the age-specific rates remained the same.

The age-specific rates in the study group are, therefore, weighted by the population in the equivalent age group of the standard population.

The formula for a directly standardised rate

(using the notation on page 4)

$$\frac{\Sigma P_i r_i}{\Sigma P_i} \text{ * constant (unit of population)}$$

Where: r_i = the rate in age group i of the study population
P_i = the population (or weight) in age group i of the standard population

Example

(using the notation and data on pages 4 and 5)

Age group	Study Group A deaths	population	rate[1]	Standard Population	Standard Population * rate
i	n_i	p_i	r_i	P_i	$P_i r_i$
<25	5	1,000	5.0	10,000	$\frac{5}{1,000}$ * 10,000 = 50
25-64	9	900	10.0	9,000	$\frac{9}{900}$ * 9,000 = 90
65+	20	500	40.0	8,000	$\frac{20}{500}$ * 8,000 = 320
Total	34	2,400	14.2	27,000	$\Sigma P_i r_i$ = 460

[1] rate per 1,000 population

In study group A, the age-standardised (or "age-adjusted" rate) per 1,000 population is:

$$\frac{\Sigma P_i r_i}{\Sigma P_i} * 1,000 = \frac{460}{27,000} * 1,000$$

$$= 17.0 \text{ per 1,000 population}$$

Note:
If study group had the same age distribution as the standard population, the rate would have been 17.0 per 1,000 population compared to the crude (or unadjusted) rate of 14.2 per 1,000 population.

Issues in the Use of Direct Standardisation

Study populations with the same age-specific rates have the same directly standardised rates

Age group	Study Group A			Study Group B		
	deaths	pop[n]	rate[1]	deaths	pop[n]	rate[1]
i	n_i	p_i	r_i	n_i	p_i	r_i
<25	5	1,000	5.0	5	1,000	5.0
25-64	9	900	10.0	10	1,000	10.0
65+	20	500	40.0	40	1,000	40.0
total	34	2,400	14.2	55	3,000	18.3

Standard Population	
Age group	pop[n]
i	P_i
<25	10,000
25-64	9,000
65+	8,000
total	27,000

Age-adjusted rates:

Study group A = 17.0/1,000 population

Study group B = 17.0/1,000 population

[1] rate per 1,000 population

Note:

The study groups have similar age-adjusted rates because they have the same age-specific rates which are weighted (ie, multiplied) by the same standard population.

Consistent inequalities between age-specific rates will produce directly standardised rates with the same inequalities

Age group i	Study Group A deaths n_i	pop^n p_i	rate[1] r_i	Study Group B deaths n_i	pop^n p_i	rate[1] r_i
<25	5	1,000	5.0	10	1,000	10.0
25-64	9	900	10.0	20	1,000	20.0
65+	20	500	40.0	80	1,000	80.0
total	34	2,400	14.2	110	3,000	36.6

Standard Population Age group i	pop^n P_i
<25	10,000
25-64	9,000
65+	8,000
total	27,000

Age-adjusted rates:

Study group A = 17.0/1,000 population

Study group B = 34.1/1,000 population

[1] rate per 1,000 population

Note:

The age-specific rates of study group B are twice those of study group A. Weighting (or multiplying) these age-specific rates by the same standard population gives an age-adjusted rate for study group B which is twice that for study group A.

The distribution of the standard population should not be markedly different from the study populations being compared

Age group	Study Group A			Study Group B		
	deaths	pop^n	rate[1]	deaths	pop^n	rate[1]
i	n_i	p_i	r_i	n_i	p_i	r_i
<25	5	1,000	5.0	5	1,000	5.0
25-64	9	900	10.0	10	1,000	10.0
65+	5	125	40.0	2	25	80.0
total	19	2,025	9.4	17	2,025	8.4

Standard Population	
Age group	pop^n
i	P_i
<25	10,000
25-64	9,000
65+	8,000
total	27,000

Age-adjusted rates:

Study group A = 17.0/1,000 population

Study group B = 28.9/1,000 population

[1] rate per 1,000 population

Note:

In contrast to the two study populations, there is a large denominator in the 65+ age group of the standard population, which places greater emphasis on this age group in which the two study groups have minor differences in rates (probably due to random variations in the small numbers involved).

Low numbers in the numerators or denominators (especially zeros) of some age groups are likely to produce unreliable age-specific rates

Age group i	Study Group A deaths n_i	pop^n p_i	rate[1] r_i	Study Group B deaths n_i	pop^n p_i	rate[1] r_i
<25	5	1,000	5.0	5	1,000	5.0
25-64	9	900	10.0	9	900	10.0
65+	20	500	40.0	1	2	500.0
total	34	2,400	14.2	15	1,902	7.9

Standard Population Age group i	pop^n P_i
<25	10,000
25-64	9,000
65+	8,000
total	27,000

Age-adjusted rates:

Study group A = 17.0/1,000 population

Study group B = 153.3/1,000 population

[1] rate per 1,000 population.

Note:

a) The study groups have similar rates in the <25 and 25 to 64 age groups.

b) In the oldest age group study group B has a much higher rate than study group A because of the small denominator (one of the two people in the age group has died). The large difference in the age-adjusted rates of the two study groups is due to this high rate.

c) Low numbers in either the numerator or denominator can have a major impact on a directly standardised rate and could lead to a spurious conclusion.

Indirect Standardisation

The basis for using indirect age-standardisation is to find out:

> *How many deaths or cases would have occurred in the study group if the age-specific rates of some standard population were applied to the population of the study group.*

The age-specific rates of some standard population are weighted by the population in the equivalent age group of the study group. This contrasts with direct standardisation where the age-specific rates of a study group are weighted by a standard population. Indirect standardisation should be used when age-specific rates in the study group(s) cannot be calculated and when there are very small denominators (populations) in the age groups of the study group.

The result of indirect standardisation is usually given as a standardised mortality ratio (SMR) if the outcome of the study is death. If the outcome is morbidity (eg, people with a condition) this index can be labelled a standardised morbidity ratio. Both indices are the ratios of the actual (or observed) number of deaths in a study population to the expected number of deaths given the standard age-specific rates. By convention 100 is used as the constant and the standard against which individual SMRs are compared. The SMR for a standard population (which provided the age-specific rates) will always be 100 as the observed and expected numbers will be equal.

The formula for calculating an SMR

(using the notation on page 4)

$$\text{SMR} = \frac{\text{total observed deaths}}{\text{total expected deaths}} * 100$$

$$= \frac{\Sigma n_i}{\Sigma p_i R_i} * 100$$

Where:

- n_i = the number of deaths or cases in age group i of the study group
- p_i = the population in age group i of the study group
- R_i = the rate in age group i of the standard group

An indirectly standardised rate can be, but rarely is, calculated by multiplying the SMR for a study population by the crude rate in the standard population.

Example (using the notation and data on pages 4 and 5)

Age group	Study Group A deaths	pop[n]	Standard population rates[1]	Expected deaths
i	n_i	p_i	R_i	$p_i R_i$
<25	5	1,000	5.0	$\frac{5}{1,000} * 1,000 = 5$
25-64	9	900	10.0	$\frac{10}{1,000} * 900 = 9$
65+	20	500	12.0	$\frac{12}{1,000} * 500 = 6$
Total	34	2,400		$\Sigma p_i R_i = 20$

[1] rate per 1,000 population

The SMR (standardised mortality ratio) for study group A is:

$$\frac{\text{total observed deaths } (\Sigma n_i)}{\text{total expected deaths } (\Sigma p_i R_i)} * 100$$

$$= \frac{34}{20} * 100$$

$$= 170$$

Age group	Study Group B deaths	pop[n]	Standard rates[1]	Expected deaths
<25	5	1,000	5.0	5
25-64	10	1,000	10.0	10
65+	40	1,000	12.0	12
Total	55	3,000		27

The SMR (standardised mortality ratio) for study group B is:

$$\frac{\text{total observed deaths}}{\text{total expected deaths}} * 100$$

$$= \frac{55}{27} * 100$$

$$= 204$$

Issues in the Use of Indirect Standardisation

An SMR for a study group should only be compared to the standard (100) and not with the SMR for another study group

Any differences (or their absence) between the SMRs of study groups may be due to differences in the age structures of the study group populations. This is because indirect standardisation weights (or multiplies) standard age-specific rates by the populations in each age group of the study group.

Age group	Study Group A			Study Group B		
	deaths	popn	rate[1]	deaths	popn	rate[1]
i	n_i	p_i	r_i	n_i	p_i	r_i
<25	5	1,000	5.0	5	1,000	5.0
25-64	9	900	10.0	10	1,000	10.0
65+	20	500	40.0	40	1,000	40.0
total	34	2,400	14.2	55	3,000	18.3

Standard Population		
Age group	popn	rate[1]
i	P_i	R_i
<25	10,000	5.0
25-64	9,000	10.0
65+	8,000	12.0
total	27,000	8.7

Direct standardisation

Age-adjusted rates:

Study group A = 17.0/1,000 population

Study group B = 17.0/1,000 population

Indirect standardisation

Standardised mortality ratios:

Study group A = 170 (20 expected deaths)

Study group B = 204 (27 expected deaths)

[1] rate per 1,000 population

Note:

a) Because the study groups have the same age-specific rates, they also have the same age-adjusted rates. In direct standardisation the age-specific rates of a study group are weighted (or multiplied) by the population in each age group of the standard population.

b) The study groups have the same age-specific and age-adjusted rates, but different SMRs. Indirect standardisation weights (ie, multiplies) the age-specific rates in the standard population by the populations in the study groups. In this example the study groups have different age structures, 42 percent of the population in study group A are under the age of 25 years compared to 33 percent in study group B.

c) Therefore, the SMR for a study group should only be compared to the standard which is usually 100. The mortality in study group A (SMR = 170) is 70 percent higher than the standard population, while the mortality in study group B (SMR = 204) is more than twice as high as the standard population.

Standardisation	*Age-specific rates*	*	*Population*
Direct standardisation	study group	*	standard
Indirect standardisation	standard	*	study group

Standard Populations and Rates

An integral aspect of standardisation is the selection of the standard population or age-specific rates. This selection is most important when direct standardisation is used as different standard populations will produce different age-adjusted rates (see page 22).

There is no unanimity about which standard populations or age-specific rates should be used, but there are three general guidelines:

a) Use the data for the entire population from which the study populations are chosen. For example, if comparing the mortality in Auckland and Wellington, or Maori and non-Maori, the total New Zealand population could be used. If the total population (eg, New Zealand) is not available, an option is to pool the study populations (ie, add the Auckland and Wellington populations). The rationale is that summing over all constituent groups should produce a population that approximates the "real" population of interest.

b) Select one of the categories of the risk factor being investigated. For example, in a study of lung cancer mortality among smokers and non-smokers, the non-smoker population could be used as the standard population.

c) Use one of the predetermined standard populations (see page 21). This is often the preferred option when calculating standard rates for the analysis of time trends or making international comparisons of disease occurrence.

Some Standard Populations Often Used in Standardisation

Age	African	World	European	Truncated World
0-	2,000	2,400	1,600	-
1 - 4	8,000	9,600	6,400	-
5 - 9	10,000	10,000	7,000	-
10 - 14	10,000	9,000	7,000	-
15 - 19	10,000	9,000	7,000	-
20 - 24	10,000	8,000	7,000	-
25 - 29	10,000	8,000	7,000	-
30 - 34	10,000	6,000	7,000	-
35 - 39	10,000	6,000	7,000	6,000
40 - 44	5,000	6,000	7,000	6,000
45 - 49	5,000	6,000	7,000	6,000
50 - 54	3,000	5,000	7,000	5,000
55 - 59	2,000	4,000	6,000	4,000
60 - 64	2,000	4,000	5,000	4,000
65 - 69	1,000	3,000	4,000	-
70 - 74	1,000	2,000	3,000	-
75 - 79	500	1,000	2,000	-
80 - 84	300	500	1,000	-
85+	200	500	1,000	-
Total	100,000	100,000	100,000	31,000

Source: Waterhouse J, Muir C, Shanmugaratnam K et al. Cancer Incidence in Five Continents IV. Lyon: IARC, Scientific Publications no. 42, 1982.

The Effect of Using Different Standard Populations

The following table of female lung cancer rates (per 100,000) shows the effect of using different standard populations in calculating directly age-standardised rates.

Area	Population Group	Standard population (see page 21) European	World	African	Truncated World
New Zealand	Maori	88.8	62.2	32.1	101.1
	Non-Maori	22.7	15.7	8.2	25.7
New South Wales		18.9	13.4	7.3	23.6
England/Wales		29.9	20.5	10.5	32.8
Hawaii	Hawaiian	55.7	39.5	22.0	76.4
Singapore	Chinese	33.0	21.8	11.2	28.0
	Malay	17.2	12.1	6.6	16.4
Canada					
Alberta		30.4	21.1	11.4	40.0
British Columbia		38.9	27.2	14.6	48.8
California					
Alameda County	White	53.7	37.9	20.1	69.7
San Francisco Bay area	White	52.4	36.9	19.4	67.4
	Black	49.7	35.9	20.5	75.9

Source: Parkin DM, Muir CS, Whelan SL, et al. Cancer Incidence in Five Continents VI. Lyon: IARC, Scientific Publications No 120, 1992.

Note: The relative risks (eg, Maori versus non-Maori) do not change much.

Issues in the Use of Standardisation

a) The decision to calculate directly or indirectly standardised rates should only be made after a careful perusal of the age-specific rates.

b) Either form of standardisation may mask aetiologically important variations in the mortality or morbidity experience in various age groups of the study group(s). For example, the age-specific rates in one study population may be higher than those in another study population at some age groups, but lower in other age groups (see *Table 2*).

c) Always refer to the age-specific rates of a study group *before* and *after* calculating any standardised rate.

Table 2: Age-Specific and Age-Adjusted Death Rates, per 100,000 Population, for White Males in Louisiana and New Mexico, 1940

Age group	Rates[1] Louisiana	New Mexico
under 1	5,693.1	13,146.8
1-4	333.3	518.1
5-14	112.4	163.0
15-24	217.8	356.7
25-34	338.7	428.1
35-44	550.4	751.5
45-54	1,389.8	1,289.1
55-64	3,029.3	2,200.9
65-74	6,048.1	4,796.5
75-84	12,602.6	10,196.5
over 85	28,959.7	22,633.7
Crude rate[1]	1,059.7	1,203.1
Age-adjusted rate[1]	13.1[2]	13.1[2]
	10.1[3]	11.7[3]

Source: Yerushalmy J. A mortality index for use in place of the age-adjusted death rate. Am J Pub Hlth 1951; 41: 907-22.

[1] rate per 100,000 population

[2] using the 1940 USA population as the standard

[3] using the 1901 England/Wales population as the standard

Reminder

Whenever standardisation is carried out, remember the advice contained in these quotations :

> *"Standardisation should never, however, substitute for a comparison of the specific rates themselves. It is these that characterise the experience (mortality, morbidity, or whatever the rate refers to) of the population being studied".*
> Source: Fleiss JL. Statistical Methods for Rates and Proportions. New York: Wiley, 1981; 239.

> *"The standardised death rate is ... a fiction ... it is not the total death-rate that actually exists in an area but the rate the area would have if, while retaining its own rates at ages, it had instead of its real population one of some particular chosen type. The fiction is useful because ... it enables summary comparisons to be made ... free from distortions which arise from age and sex differences in the existing populations. The object throughout is, therefore, comparison; a standardised death rate alone has no meaning".*
> Source: Hill AB. Principles of Medical Statistics. London: The Lancet 1971; 158.

References

Further information about standardisation can be obtained from the following references:

Introductory

Ahlbom, A, Norell, S. Introduction to Modern Epidemiology. Chestnut Hill, Massachusetts: Epidemiology Resources, 1984.

Barker, DJP, Rose, G. Epidemiology in Medical Practice. Edinburgh: Churchill Livingstone, 1976.

Friedman, GD. Primer of Epidemiology. New York: McGraw-Hill, 1980.

Rose, G, Barker, DJP. Epidemiology for the Uninitiated. London: British Medical Journal, 1979.

Advanced

Armitage, P. Statistical Methods in Medical Research. Oxford: Blackwell, 1974.

Fleiss, JL. Statistical Methods for Rates and Proportions. New York: Wiley, 1981.

Kleinbaum, DG, Kupper, LL, Morgenstern, H. Epidemiologic Research: principles and quantitative methods. Belmont, California: Lifetime Learning Publications, 1982.

MacMahon, B, Pugh, TF. Epidemiology: principles and methods. Boston: Little Brown, 1970.

Rothman, KJ. Modern Epidemiology. Boston: Little Brown, 1986.

Exercises

Exercise 1

a) What is the numerator of a rate?

b) What is the denominator of a rate?

c) What is a crude rate?

d) What is an age-specific rate?

Exercise 2

a) What does age-standardisation attempt to do?

b) What process is involved in age-standardisation?

Exercise 3

a) What age-specific rates and populations are used in direct standardisation?

b) What age-specific rates and populations are used in indirect standardisation?

Exercise 4

a) Provide two instances when direct standardisation should be used with caution?

b) Indirect standardisation usually produces an index labelled a SMR. What is this?

c) When would indirect standardisation be used in preference to direct standardisation?

d) Why should the SMR of a study group only be compared to the SMR of the standard population rather than the SMR of another study group?

Exercise 5

Using the following data:

a) Calculate the crude rates and age-specific rates per 100,000 population for the two study groups.

b) Using the truncated world standard population (see page 21), calculate the directly age-standardised mortality rates per 100,000 population for the two study groups.

Age group i	Study Group A deaths n_i	Study Group A pop^n p_i	Study Group B deaths n_i	Study Group B pop^n p_i
35-44	4	14,000	12	165,000
45-54	20	16,000	80	160,000
55-64	50	11,000	180	125,000
total	74	41,000	272	450,000

Exercise 6

Using the following data and direct standardisation:

a) Calculate the crude rates and age-specific rates per 1,000 population for study group A and study group B.

b) Calculate the age-adjusted rates per 1,000 population.

Age group i	Study Group A deaths n_i	Study Group A pop^n p_i	Study Group B deaths n_i	Study Group B pop^n p_i
< 25	80	10,000	250	25,000
25-44	165	15,000	180	15,000
45-64	375	25,000	160	10,000
total	620	50,000	590	50,000

Standard Population Age group i	pop^n P_i
<25	45,000
25-64	35,000
65+	20,000
total	100,000

Exercise 7

Using the following data and indirect standardisation:

a) Calculate the expected number of deaths in the two study groups.

b) Calculate the SMRs for the two study groups.

Age group i	Study Group A deaths n_i	Study Group A pop^n p_i	Study Group B deaths n_i	Study Group B pop^n p_i
< 25		10,000		25,000
25-64		15,000		15,000
65+		25,000		10,000
total	620	50,000	590	50,000

Standard Population Age group i	rate[1] R_i
<25	9.4
25-64	11.5
65+	15.3

[1] rate per 1,000 population

Exercise 8 Using the following data:

a) For each study group, calculate the age-specific death rates per 1,000 population and the crude (unadjusted) death rate per 1,000 population. How do the rates in the two study groups compare? Explain any discrepancy.

b) Calculate directly age-standardised mortality rates for study groups A and B.

c) Use indirect standardisation to calculate the SMR for each study group.

d) Explain any difference between the directly age-standardised rates and the SMRs.

Age group	Study Group A		Study Group B	
	deaths	pop^n	deaths	pop^n
i	r_i	p_i	r_i	p_i
< 25	6	1,500	8	2,000
25-44	32	2,000	40	2,500
45-64	30	1,500	10	500
total	68	5,000	58	5,000

Age group	Standard Population pop^n	rate[1]
i	P_i	R_i
<25	45,000	6.0
25-64	35,000	10.0
65+	20,000	25.0
total	100,000	

[1] rate per 1,000 population

Answers to the Exercises

Exercise 1

a) The numerator is the total number of deaths or cases used in the calculation of a rate.

b) The denominator of a rate is the total population-at-risk or person-years-at-risk. The numerator should always be contained in the denominator. For example, a female death rate from lung cancer will only include females in the denominator, while the male rate of prostate cancer will only include males in the denominator.

c) A crude rate is the actual disease or death rate observed in a study group. It is a count of all deaths or the numbers of people with a particular disease or condition (the numerator) divided by the number of people in that population who were at risk of death or the disease (the denominator), expressed as a unit of the population. A crude rate takes no account of the age distribution of the population-at-risk, the person-years-at-risk, or any effect these may have on the resulting rate.

d) An age-specific rate is the actual rate of death or disease in a particular age group (eg, 15 to 19 years of age, 25 to 44 years of age).

Exercise 2

a) Age standardisation is a method which adjusts or controls for the effects of differences between the age structures of populations.

b) Age standardisation involves weighting (or multiplying) age-specific rates by the population-at-risk or person-years-at-risk in a specific age group.

Exercise 3

a) Direct standardisation weights (or multiplies) the age-specific rates in a study group by the population in each age group of the standard population.

b) Indirect standardisation weights (or multiplies) the age-specific rates in a standard population by the population in each age group of the study group.

Exercise 4

a) Direct standardisation should be used with caution when:
   (1) the age structure of the standard population is markedly different from the age structure of the study groups; and
   (2) there are low numbers of deaths or cases, especially zeros, in the numerator or denominator.

b) SMR means a standardised mortality or morbidity ratio.

c) Indirect standardisation is preferred to direct standardisation when:
   (1) the age-specific rates in the study groups are likely to be "unstable" because of low numbers in the denominator; and
   (2) the age-specific rates cannot be calculated for the study groups.

d) In calculating a SMR, the age-specific rates in the standard population are weighted (or multiplied) by the populations-at-risk in each age group of the study groups. Any difference or the absence of any difference between the SMRs of the study groups may, therefore, result from the study groups having different age structures. A SMR for a study group should only be compared to the standard (ie, 100).

Exercise 5

a) The crude rates: study group A: 180.5 per 100,000 population; study group B: 60.4 per 100,000 population.

The age-specific rates per 100,000: study group A: 28.6, 125.0, and 454.5; study group B: 7.3, 50.0, and 144.0.

b) The directly age-standardised rates are: study group A, 172.7 per 100,000 population; study group B, 57.7 per 100,000 popuation.

Exercise 6

a) Crude rates: study group A: 12.4 per 1,000 population; study group B: 11.8 per 1,000 population. Age-specific rates per 1,000 population: study group A: 8.0, 11.0, and 15.0; study group B: 10.0, 12.0, and 16.0.

Age-adjusted rates: study group A: 10.5 per 1,000 population; study group B: 11.9 per 1,000 population.

Exercise 7

a) Standardised mortality ratios (SMRs) for the two study groups are: study group A: 95.5 (expected number of 649.0); study group B: 105.3 (expected number 560.5).

Exercise 8

a) Age-specific rates per 1,000 population are: study group A: 4.0, 16.0, and 20.0; study group B: 4.0,16.0, and 20.0. The crude rates per 1,000 population are: study group A, 13.6 per 1,000; study group B, 11.6 per 1,000.

The age-specific rates for the two populations are the same, but the crude rates are quite different. This is because study group A has a higher proportion of its population (30%) in the older age group compared to study group B (10%). Hence, the crude rate for study group A puts more weight on the older age-group where the death rate is higher.

b) The directly age-standardised rates are the same (11.4 per 1,000 population) for both study groups. This is expected as the two study groups have the same age-specific rates.

c) The SMR for study group A is 102.3 (66.5 expected deaths); and the SMR for study group B is 117.2 (49.5 expected deaths).

d) The directly age standardised rates are similar, but the SMRs are different because of the differences in the age distributions of the two study groups. SMRs for study groups are not really comparable and comparing them can lead to spurious findings. Study group SMRs should only be compared to the standard (100).