2 Data methodology and interpretation

This chapter discusses key analytical concepts used to compare mortality rates across regions, issues related to the quality and interpretation of the data, and identifies areas of work requiring further improvement and development. This chapter is intended to provide clear guidance about interpreting the findings presented in chapters 3–9.

2.1 Methodology

Several analytical concepts have been used to compare mortality rates across the regions. There are two in particular that are crucial to understanding the discussion that follows. While these and other concepts are explained more fully in Appendix A, they are, briefly:

- **standardised mortality ratio (SMR)**—the ratio of the actual number of deaths in an area to the number expected if Major Cities death rates for the relevant group had applied in that area
- **'excess' deaths**—the difference between the actual number of deaths in an area and the number that would have occurred if Major Cities death rates had applied.

Indirect age standardisation involved the use of:

- age-specific death rates for people living in Major Cities as the standard for calculating the expected numbers of deaths in each area
- age-specific death rates for non-Indigenous Australians living in Major Cities as the standard for calculating the expected numbers of deaths of non-Indigenous Australians in each area and of Indigenous Australians in Queensland, Western Australia, South Australia and the Northern Territory.

'Excess' deaths have been reported because although SMRs provide a measure of inequity, they do not provide a measure of magnitude (that is, an understanding of the absolute size of disadvantage for particular causes of death in each region in terms of human lives lost).

Reporting of mortality trends relies on previously published analyses (AIHW 2006a). It compares the numbers of deaths in each year 1992 to 2003, with the number that would be expected if Major Cities age-specific rates in the period 2001–03 were applied to the populations in each area in those years. The trend is estimated using weighted least squares.

Mortality for the period 2002–04 is compared with that for 1997–99, both in terms of whether there have been real (absolute) changes in mortality between the two periods, and also in terms of whether there have been changes to the death rates in regional areas relative to those in Major Cities at the time.

In order to assess absolute changes between the two periods (for example, whether Inner Regional rates in 2002–04 were higher compared with what they were in 1997–99), the number of deaths in each area in each period were compared to the number that would have occurred if 2002–04 Major Cities age-specific rates were applied to the populations in each area in each period. In essence this answers the question 'have death rates increased or declined?'.

To assess changes in the relative difference between areas (for example, whether, death rates in Inner Regional areas have remained similar to those in contemporary Major Cities,

whether they have increased relative to those in contemporary Major Cities or whether they have decreased relative to those in contemporary Major Cities), the number of deaths in each area in each period were compared to the number that would have occurred if contemporary Major Cities age-specific rates were applied to the populations in each area in each period. In essence this answers the question 'has the size of the gap between Major Cities and regional/remote areas changed?'.

Reporting of death rates in coastal and inland regional and remote areas in 2001 has used Major Cities age-specific rates in that year as the standard.

Age-specific death rates have been reported throughout the report because summary measures like SMRs can sometimes mask important patterns.

2.2 Indigenous mortality

Previous descriptions of mortality, and other measures of health, have shown poorer outcomes in more remote areas (AIHW: Strong et al. 1998), but it is possible that a lot of this difference is a result of poor Indigenous Australian health. To assess whether the poorer health in more remote areas reflects the influence of remoteness or Indigenous Australian health, mortality for the Indigenous Australian and non-Indigenous Australian populations should be compared with mortality for the total population. However, two issues affect the reporting of data for Indigenous Australians:

- Concerns about the inter-regional differences in the accuracy of the recording of Indigenous Australian deaths prevent reporting on Indigenous Australian mortality separately for the five regions used in this report. If this analysis was completed, any differences between areas may have reflected accuracy of the records rather than real differences in mortality. Consequently, overall, rather than regional, mortality rates for Indigenous Australians are presented.
- Identification of Indigenous Australian mortality was considered to be most reliable in Queensland, Western Australia, South Australia and the Northern Territory during the study period. Overall mortality rates for Indigenous Australians have been calculated using data from these jurisdictions only.

Because a 'non-Indigenous Australian' person has been defined in this report as someone who is not identified as Indigenous Australian, under-identification of Indigenous Australians will necessarily mean over-reporting of non-Indigenous Australians in the mortality data.

For many of the causes of death examined, rates for Indigenous Australians are much higher than for non-Indigenous Australians from any area. Elevated death rates in remote areas may be a consequence of the proportionally large number of Indigenous Australians in those areas, and high overall Indigenous Australian mortality.

Frequently, death rates for elderly non-Indigenous Australians from remote areas appear substantially lower than for their Major Cities counterparts, while rates for younger people from remote areas are higher than for those in Major Cities. It is possible that this effect may be due to elderly people in poorer health migrating to less remote areas where they can access services, leaving the healthier individuals, who have lower death rates. To control for this apparent effect, death rates for some populations younger than 65 years have been presented alongside those for the total population.

2.3 Notes on reading and interpreting the main tables in each chapter

Each set of two tables describes:

- in the first table, death rates for the total population
- in the second table, death rates for the non-Indigenous Australian population in each Remoteness Area and for Indigenous Australians in Queensland, Western Australia, South Australia and the Northern Territory.

The structure of the two tables is similar.

Three sets of columns across the page report for males, females and for persons.

Within each set in the first table (reporting mortality for all people) are five columns which provide details for MC, IR, OR, R and VR areas. In the second of the tables (relating to Indigenous Australian and non-Indigenous Australian mortality) there are six columns each for males, females and persons—the five regional columns for non-Indigenous Australian mortality, and a single column for Indigenous Australian mortality in Queensland, Western Australia, South Australia and the Northern Territory.

The top half of the table

The top half of each table reports death rates, usually as standardised mortality ratios (SMR), ratios which compare the number of deaths in a population with the number that would be expected if age-and sex-specific rates of death in a specified standard population were to apply to the population in each area. For example, if there were 100 deaths in a population, but only 50 expected, then the ratio would be 2.00, and we could say that the death rate in the population was twice that of the standard population. SMRs have the advantage of being useful in comparing death rates in small populations or for causes of death that are relatively uncommon. However, their disadvantages are that the comparison of death rate is, strictly speaking, with the standard population, and that the ratio does not give a measure of the burden due to that cause of death (for example, the SMR for disease X may be 5.5 (and therefore alarming), while for disease Y it may be 1.1 (and therefore less alarming), however disease X may be very rare, killing one person per year, while disease Y may be common, killing 1,000 people per year).

The first column for males, females and persons contains crude death rates for the Major Cities population in 2002–04 (expressed as deaths per 100,000 population). These are provided because, by definition, the SMR for Major Cities is equal to 1.0 in every case (so therefore there is no point reporting it) and because a crude rate, like a count of the number of deaths, provides a measure of the burden of mortality; for example, a crude death rate of 5 per 100,000 population per annum indicates less of a burden than a crude death rate of 2,000 per 100,000 per annum.

Death rates (crude death rates and SMRs) are reported in some detail for the three-year period 2002–04, for each life stage age group, for the total population and for the population younger than 65 years. The SMRs presented here compare the actual number of deaths in each population with the number expected if the age-and sex-specific death rates in Major Cities in 2002–04 had applied to these populations.

A little lower down the table, death rates for the previous reporting period (1997–99) are detailed. The first two rows (shaded) in this section use Major Cities age and sex specific rates in 1997–99 as the standard and compare death rates in each of the areas with that in Major Cities in the same year (1997–99). Consequently, these first two shaded rows can be used to compare regional and remote death rates with those in Major Cities within 1997–99.

The second two (unshaded) rows (marked with a '†') use Major Cities age- and sex-specific rates in 2002–04 as the standard and compare death rates in each of the areas (including Major Cities) in 1997–99 with death rates in Major Cities in 2002–04. These second two (unshaded) rows can be used to compare death rates in each of the areas (including Major Cities) directly with death rates in Major Cities in 2002–04 (and indirectly with death rates in each of the areas in 2002–04).

For example, with reference to Table 5.10 (coronary heart disease (CHD)):

- In 1997–99, CHD death rates in Very Remote areas were 1.4 times those in Major Cities at the time. Several years later in 2002–04, CHD death rates in Very Remote areas were still 1.4 times those in Major Cities at that time.
- In 1997–99 CHD death rates in Major Cities were 1.3 times what they were to become in 2002–04 (that is, rates in Major Cities declined substantially between these two periods). In Very Remote areas between 1997–99 and 2002–04, the CHD death rate had declined from 1.9 to 1.4 times the 2002–04 Major City rate.
- Death rates due to CHD have declined in all areas, but the death rate in Very Remote areas due to this cause is still 1.4 times higher than in Major Cities (because rates in Major Cities declined at approximately the same rate as those in Very Remote areas).

SMRs for males, females and persons cannot be compared with one another as they relate to different standards. Similarly, SMRs cannot be compared across age groups for the same reason and comparisons between SMRs for different causes of death cannot validly be made.

The bottom half of the table

The bottom half of the table describes the actual number of deaths and 'excess' deaths that occurred in each population. These provide an understanding of where most of the burden falls, both in terms of the actual number of deaths and 'excess' deaths. 'Excess' deaths are deaths in excess if what would be expected if Major Cities rates had applied in each population.

The first seven rows detail the number of 'excess' deaths in each age group in each area in 2002–04; this provides an understanding of which age groups and genders contribute most towards higher death rates in each area.

The next four rows detail the total number of 'excess' deaths, the total number of deaths, the total number of 'excess' deaths for those younger than 65 years and the total number of deaths of people younger than 65 years, annually, in 2002–04. For example, annually in Very Remote areas, there were 149 deaths due to CHD in 2002–04 (see Table 5.10), of which 45 were 'excess' deaths (in other words, 104 deaths were expected, but 149 occurred on average each year). Of the 75 deaths due to CHD that occurred amongst people younger than 65 years in Very Remote areas, 53 were 'excess' deaths (in other words, 22 deaths were expected, but 75 occurred). 'Excess' deaths in Very Remote areas were mainly among 25–64 year olds, with fewer deaths than expected amongst those 75 years and older.

The last six rows of the table relate to deaths and 'excess' deaths in 1997–99. The first three of these six relate to 'all ages', while the last three of these six relate to people younger than 65 years.

The number of 'excess' deaths has been calculated in two ways:

- Shaded rows 1 and 4 have used 1997–99 Major Cities rates of death as the basis for calculating the number of 'excess' deaths. For example, in Very Remote areas in 1997–99, there were 169 deaths due to CHD, of which 51 were in excess of what would have been expected if Major Cities rates at the time had applied in Very Remote areas. Of these deaths, 84 were of people younger than 65 years, with 59 of these being in excess of what would have been expected if Major Cities rates for the period had applied (that is, 25 expected deaths, with 84 observed).
- Unshaded rows 2 and 5 (marked with a '†') have used 2002–04 Major Cities rates of death as the basis for calculating the number of 'excess' deaths in 1997–99. These numbers of 'excess' deaths are directly comparable to the number of 'excess' deaths calculated for 2002–04. For example, in Very Remote areas in 1997–99, 80 of the 169 deaths due to CHD were in excess of what would have been expected if 2002–04 rates of CHD death had applied. This compares with 45 excess deaths out of 149 deaths due to CHD in 2002–04. The advantage of this form of analysis is that the excess in both periods is based on one standard, and it is clear that the absolute number of deaths has declined by 20 per annum, and the number of 'excess' deaths has declined by 35 (from 80 to 45) per annum.

SMRs that are statistically significantly different from 1.0 (that is, different from Major Cities) are in bold print and accompanied by an asterisk.

The data in the tables are from AIHW analysis of the AIHW mortality database.

Technical notes on data presentation

- Percentages or numbers in tables may not add to 100 or the total due to rounding.
- ICD-10 (International Classification of Diseases, 10th revision) codes for the described causes of death are listed in Appendix B.
- All standardisation of death rates has been indirect using Major Cities rates for males and females for the period 2002–04 (2001 for the coastal analysis), or Major Cities rates for non-Indigenous Australian males and females for the period 2002–04 (2001 for the coastal analysis). The former have been used to standardise rates for the total (Indigenous Australian plus non-Indigenous Australian) population, while the latter have been used to standardise rates for Indigenous Australian and non-Indigenous Australian populations separately.
- In this report, names of specific areas defined by the ASGC Remoteness Areas have been capitalised (for example, Inner Regional, Remote, and Very Remote). Where Inner Regional and Outer Regional areas are taken together, they are referred to as 'regional', when Remote and Very Remote areas are taken together, they are referred to as 'remote'.
- 'Excess' deaths are calculated by subtracting the expected number of deaths from the number observed. Expected deaths are the number of deaths expected annually if death rates found in Major Cities are applied to the populations living in each of the other areas. 'Excess' deaths provide an indication of the extra burden of mortality in each area.

- Where there were fewer deaths than expected, this report states either (for example) fewer deaths than expected annually, or –5 'excess' deaths annually; both expressions mean the same thing.
- All statements about rates of death in this report are based on the ratio of observed to expected deaths. If there are twice as many deaths as expected, then the rate of death can be assumed to be twice that of the comparison population.
- Where rates are statistically significantly different from one another, they are referred to in the text as significantly different. Statistical significance is at the 95% level.
- In the text, where reference is made to 'Major Cities, Inner Regional, Outer Regional, Remote and Very Remote areas', the term 'the five areas' has been used. Where there is reference to 'Inner Regional, Outer Regional, Remote and Very Remote areas', the term 'the four areas outside Major Cities' has been used.
- Graphs are presented as bar charts with error bars (for example, Figure 3.1). These error bars indicate the values of the lower and upper 95% confidence levels. We can be 95% sure that, if the underlying rates remained the same and we calculated the death rate in the preceding year or the next year, the calculated rate would lie somewhere between the two presented error bars. In the graph, the top of the column (between the two error bars) indicates our best point estimate with the available data. There is one chance in 20 that the true value lies outside the error bars. Error bars do not provide any indication of the level of uncertainty due to bias in the data (for example, potential bias as a result of different accuracy in the identification of Indigenous Australian deaths in each area). Columns representing estimates of SMRs for non-Indigenous Australians from Remote and Very Remote areas have a dashed outline, indicating uncertainty about identification issues discussed in Appendix A.
- Statistically significant numbers are indicated in bold and with an asterisk.

2.4 Further developments

This section outlines several areas of this work which require further improvement or development.

- Improvements in the identification of Indigenous Australians in the mortality data collection are crucial to being able to describe differences in mortality across remoteness in the future.
- Estimates of the accuracy of Indigenous Australian identification in each area are critical for the utilisation of current and historical mortality data to assess differences in mortality for Indigenous Australians in each area.
- Descriptions of regional mortality rely on relatively crude allocation of regional category on the basis of Statistical Local Area (SLA), because the boundaries of SLAs and remoteness categories seldom coincide exactly. Geocoding of residential location would allow more precise allocation, and would also facilitate more powerful epidemiological work (for example, identification of disease clusters); however, a move to geocoding would need to incorporate substantial confidentiality safeguards.
- The lower death rates of the aged in remote areas may be due to the migration of the frail aged to less remote areas where they can access services—although little has been published to support this hypothesis. Further work in this field is recommended.

- This report does not include information on deaths specifically due to occupational accidents (although these will be included amongst motor vehicle accidents, other injuries and so on). This is an important issue that requires further investigation.
- The effect of income and education on regional differences in mortality has not been explored in this report. It is possible that these factors would explain some of the regional differences in mortality.
- Although this report describes, amongst others, high death rates due to ischaemic heart disease, 'other' circulatory disease, and motor vehicle accidents, it is not clear whether these differences are due to higher overall risk, or due to lower levels of access to health services or both. Further work to identify the contribution of risk and access would be useful.
- Further refinement of regional analysis to incorporate coastality may be useful in further refining understanding of inter-regional differences and may lead to more constructive allocation of resources and management of rural health issues.
- Further inland/coastal work would benefit from:
  - development of population data to allow analysis over a wide time period and to provide well developed counts of Indigenous Australians and non-Indigenous Australians in each area
  - development of coastality/remoteness concordances for years other than 2001
  - calculation of inland/coastal/remoteness results for years other than 2001
  - calculation of death rates for specific causes of death, so as to better understand the possible cause of the inland/coastal differences
  - building an Indigenous Australian component into the analysis to asses whether the differences disappear or change appreciably if non-Indigenous Australian mortality only is considered
  - assessing the relative importance of a range of factors in predicting higher death rates. This would require some sort of regression or analysis of variance to assess the importance of a range of factors: remoteness, coastality, SEIFA, Indigenous status, access to services, SLA population growth, and so on.