

Appendix C: Data sources

To provide a comprehensive picture of national breast cancer statistics in this report, a range of data sources were used, including AIHW and external data sources. These data sources are described in this appendix.

Australian Cancer Database

The Australian Cancer Database (ACD) is a database that holds information about 1.8 million cancer cases of Australians who were diagnosed with cancer (other than basal cell and squamous cell carcinomas of the skin) between 1982 and 2006. Data from this source are used in a number of chapters including Chapters 2, 4 and 5.

Each Australian state and territory has legislation that makes the reporting of all cancers (excluding basal cell and squamous cell carcinomas of the skin) mandatory. Pathology laboratories and Registrars of Births, Deaths and Marriages across Australia must report on cancer cases, while in some (but not all) jurisdictions, hospitals, radiation oncology units and nursing homes must also report on cancer cases.

The data from the pathology laboratories, Registrars and other sources (as applicable) are sent to state and territory cancer registries. On an annual basis, these registries provide data to the AIHW. The AIHW compiles and maintains the ACD, in partnership with the Australasian Association of Cancer Registries.

The data provided to the AIHW by the states and territories include, at a minimum, an agreed set of items that provide information about the individual with the cancer, the characteristics of the cancer and, where relevant, deaths from malignant tumours (see Table C.1). In addition to the agreed set of items, registries often provide other data which are also included in the ACD. For example, data on ductal carcinoma in situ (DCIS) are not part of the agreed ACD data set but are regularly provided by the state and territory registries.

Once the data are received from the state and territory cancer registries, the AIHW assembles the data into the ACD. Internal linking checks are undertaken to identify those who had tumours diagnosed in more than one state or territory; this process reduces the degree of duplication within the ACD to a negligible rate. The ACD is also linked with information on deaths (from the National Death Index) in order to add information on which people with cancer have died (from any cause). Any conflicting information and other issues with the cancer data are resolved through consultation with the relevant state or territory cancer registry.

The registration of cases of cancer is a dynamic process such that records may be modified if new information is received. Thus, records in the cancer registries are always open and they are updated as required. In order for these changes to be incorporated into the ACD, a new complete file for all years of cancer data is provided by each of the jurisdictions annually. As a result, the number of cancer cases reported by the AIHW for any particular year may change slightly over time and, in addition, data published by a cancer registry at a certain point in time may differ to some extent from what is published by the AIHW.

Table C.1: Agreed set of items to be provided by the states and territories to the AIHW for inclusion in the Australian Cancer Database

Person-level attributes	Tumour-level attributes
Person identification number (assigned by the state/territory)	Tumour identification number (assigned by the state/territory)
Surname	Date of diagnosis
First given name	Date of diagnosis flag
Second given name	Age at diagnosis
Third given name	ICD-O-3 ^(a) topography code
Sex	ICD-O-3 ^(a) morphology code
Date of birth	ICD-10 ^(b) disease code
Date of birth flag	Most valid basis of diagnosis
Indigenous status	Statistical local area at diagnosis
Country of birth	Postcode at diagnosis
Date of death	Melanoma thickness (mm)
Age at death	
Cause of death	

(a) International Classification of Diseases for Oncology, 3rd edition.

(b) International Statistical Classification of Diseases and Related Health Problems, 10th revision.

Source: AIHW 2009c.

Non-melanoma skin cancers

Data on all types of cancer, other than two types of non-melanoma skin cancer (NMSC), are reportable and collected by the state and territory registries. The two most common types of NMSC – namely, basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) – are not reportable and are thus not generally recorded in cancer registries in Australia. These two types of skin cancers are by far the most frequently diagnosed cancers in Australia for both males and females (AIHW & CA 2008). A number of other, rarer types of cancer also fall within the NMSC category (e.g. Merkel cell lesions, Kaposi sarcoma and cutaneous lymphoma) and these are reportable cancers.

In the past, the agreed approach was to exclude all NMSC cases from the incidence data produced by the AIHW. However, for the first time this year, a new approach is used whereby all cases that pertain to reportable forms of NMSC are included in the data; as previously, no data on BCC and SCC are included. To implement this change, the state and territory registries were asked to supply, along with the usual data, information on all NMSC cases other than BCC and SCC for 2006 and for all previous years, where possible. All of the registries were able to provide such data for 2001 to 2006, with only some being able to provide data for earlier years. Thus the data on non-melanoma skin cancers other than BCC and SCC may be incomplete before 2001.

BreastScreen Australia Program data

Data from the BreastScreen Australia Program were used in Chapter 7 to indicate the number of women who had a screening mammogram through the BreastScreen Australia Program.

These data are supplied annually to the AIHW by individual state and territory BreastScreen programs for monitoring purposes. They are compiled by the AIHW and reports are produced annually (AIHW 2009b).

Disease Expenditure Database

Expenditure data are used in Chapter 9 of this report to describe health expenditure on breast cancer. These data were obtained from the Disease Expenditure Database which is maintained by the AIHW.

Since 1998, the AIHW has had responsibility for developing estimates of national health expenditure. Data for this purpose are obtained from a wide variety of sources in the public and private sectors, with the Australian Bureau of Statistics, the Australian Government Department of Health and Ageing, and state and territory health authorities providing most of the basic data. Other major sources are the Department of Veterans' Affairs, the Private Health Insurance Administration Council, Comcare, and the major workers compensation and compulsory third-party motor vehicle insurers in each state and territory.

The definition of 'all cancers' used in Chapter 9 is somewhat different from that used in earlier chapters as it only includes ICD-10 'C' codes and excludes those malignant cancers with ICD-10 'D' codes (such as polycythaemia vera). Separate expenditure data were not readily available for the required set of ICD-10 'D' cancers. Since the forms of malignant cancers covered by the ICD-10 'D' codes are not common (see AIHW & AACR 2008), their exclusion is not expected to have a large effect on the health expenditure estimates shown in this report.

Further information about the Disease Expenditure Database can be found in the annual health expenditure reports published by the AIHW (AIHW 2008c).

Ductal carcinoma in situ data

Data on the number of cases of ductal carcinoma in situ (DCIS) in females are provided in Chapter 2. These data relate to the period from 1995 to 2005. The first year in which national data are available is 1997 since the 1996 Northern Territory data on the number of DCIS cases was of uncertain quality, as was the 1995 and 1996 data for South Australia.

Cancer registries in Australia generally apply a '4-month' rule to DCIS data. That is, if a woman is diagnosed with invasive breast cancer in the same site of the breast within 4 months of being diagnosed with DCIS, it is assumed that she had invasive breast cancer at the time of the DCIS diagnosis. In such cases, the DCIS case is deleted and recoded as invasive breast cancer by the cancer registries.

The DCIS data presented in Chapter 2 were provided to the AIHW by the state and territory cancer registries. Checks on the data were completed at the AIHW to remove duplicate cases. In addition, AIHW checked for inconsistencies in the application of the '4-month' rule. When

an invasive breast cancer was recorded for the same woman within 4 months of the DCIS being diagnosed, the AIHW removed that DCIS case from the database. This led to the removal of 125 DCIS cases, or less than 1% of DCIS cases, from the database. Ideally, the 4-month rule would only have been applied by the AIHW in cases where the invasive breast cancer was recorded as being in the same anatomical site of the breast as the DCIS, but this information was not available and thus could not be taken into account.

Medicare Australia data

Data from Medicare Australia were used in Chapter 7 to provide information on the annual number of mammograms subsidised under the Medicare Benefits Schedule (MBS). These data are available from the Medicare Australia website at:

www.medicareaustralia.gov.au/statistics/mbs_item.shtml.

Medicare Australia's statistics are based on the items and groups in the MBS and can be broken down by gender and age group. As mammography reimbursement is available for both females and males, data were provided for both sexes in this report. The relevant MBS item numbers are 59300 (mammography of both breasts) and 59303 (mammography of one breast).

The Medicare Australia statistics include data on only those services that were performed by a registered provider for services that qualify for a Medicare Benefit and for which a claim had been processed by Medicare Australia. They do not include services provided by hospital doctors to public patients in public hospitals or services that qualify for a benefit under the Department of Veterans' Affairs National Treatment Account.

The month and year of the procedure is determined by the date on which the service was processed by Medicare Australia, not the date on which the service was provided.

National Death Index

Cancer incidence data were linked to the National Death Index (NDI) in order to provide survival and prevalence information (Chapters 4 and 5). The NDI is a database that is maintained by the AIHW; it contains information on all deaths that have occurred in Australia since 1980.

The NDI database comprises the following variables for each deceased person: name; alternative names (including maiden names); date of birth (or estimated year of birth); age at death; sex; date of death; marital status; Indigenous status; state or territory of registration; and registration number. Cause of death information in a coded form is also available. For records to 1996, only the code for underlying cause of death is available. For records from 1997, the codes for the underlying cause of death and all other causes of death mentioned on the death certificate are available.

This database exists solely for research linkage purposes such as to gain epidemiological mortality information on individuals in a particular cohort, or with a known disease state. Ethics approval is required for the NDI to be utilised for any particular research project.

National Hospital Morbidity Database

Data from the National Hospital Morbidity Database (NHMD) are used in Chapter 8 of this report to examine the number of breast cancer-related hospitalisations. The NHMD contains demographic, diagnostic, procedural and duration of stay information on episodes of care for patients admitted to hospital. This annual collection is compiled and maintained by the AIHW, using data supplied by state and territory health authorities. Information from almost all hospitals in Australia is included in the database: public acute and public psychiatric hospitals, private acute and psychiatric hospitals, and private free-standing day hospital facilities. The database is episode-based and it is not possible to count patients individually.

Data are held for the years from 1993–94 to 2007–08. However, around 1998–99, hospitals across Australia began to implement a change in the classification system used to code the diagnosis for hospitalisations (i.e. from ICD-9 to ICD-10). The first full year for which national data are available using ICD-10 is 1999–00. Hence, in Chapter 8, only those data that were coded using ICD-10 are presented.

Note that the hospitalisations data presented in this report exclude those hospitalisations for which the care type was reported as *newborn*, *hospital boarder* or *posthumous organ procurement*. Thus, it includes all other admitted care hospitalisations including those with a care type of *acute care*, *rehabilitation care* and *palliative care*.

Comprehensive hospital statistics from this database are released by the AIHW on an annual basis (AIHW 2009a). Further information about this data source is available in those annual reports.

National Mortality Database

Data from the National Mortality Database are used in Chapter 3 of this report to provide statistical information on mortality in Australia due to breast cancer.

The registration of deaths has been compulsory since the mid-1850s and this information is registered with the relevant state and territory Registrar of Births, Deaths and Marriages. Since 1906, the Commonwealth Statistician has compiled the information collected by the Registrars and published national death information.

The National Mortality Database, which is maintained by the AIHW, currently contains information for all deaths in Australia registered from 1964 to 2006. Data back to 1906 are also available for a number of causes of death.

Information is provided to the AIHW by the Registrars of Births, Deaths and Marriages and coded nationally by the Australian Bureau of Statistics (ABS). Death certificate information is standardised and coded according to rules set forward in various versions of the ICD. The deaths have been coded to reflect the underlying cause of death. As well, in recent years, multiple causes of death have been added to the mortality data.

Over time, changes have been made to the coding and processing of mortality data that affect comparability of the data. For instance, data for holdings for 1987 to 1996 were manually coded using the ninth revision of the ICD. Data holdings for 1997 onwards were coded using ICD-10, using an automated system with slightly different coding rules. The change to the coding and processing of mortality data introduced a break in the time series. The ABS has developed comparability factors so that a time series may still be derived (ABS

2009d). As noted in Appendix A, for breast cancer, the comparability factor for ICD-9 to ICD-10 is 0.98.

In the National Mortality Database, the number of deaths is based on the 'year of death', except for the most recent year, where 'year of registration of death' is used. While for the most part, year of death and its registration coincide, deaths at the end of each calendar year may be held over until the following year, as will deaths whose cause requires further examination by a coroner (e.g. possible suicides).

Population data

Throughout this report, population data were used to derive rates of, for example, cancer incidence and mortality. The population data were sourced from the ABS Demography section using the most up-to-date estimates available at the time of analysis.

To derive their estimates of the resident populations, the ABS uses the 5-yearly Census of Population and Housing data and adjusts it as follows:

- all respondents in the Census are placed in their state or territory, statistical local area and postcode of usual residence; overseas visitors are excluded
- an adjustment is made for persons missed in the Census (approximately 2%)
- Australians temporarily overseas on Census night are added to the usual residence Census count.

Estimated resident populations are then updated each year from the census data using indicators of population change, such as births, deaths and net migration. More information is available from the ABS website <www.abs.gov.au>.