# Contents

1. Overview 1
2. Introduction and Purpose 2
3. Data Gaps 3
4. Data Sets 5

| National Cancer Statistics Clearing House (AIHW) | 5 |
| Clinical cancer registries (states/territories/professional colleges/health agencies) | 9 |
| National mortality database (ABS/AIHW) | 13 |
| National Death Index (AIHW) | 16 |
| National hospital morbidity database (AIHW) | 18 |
| Hospital case-mix data for privately insured care (DOHA) | 24 |
| Private hospital data bureau (DOHA) | 28 |
| Hospital outpatient minimum data set (AIHW) | 31 |
| Medicare Australia (DOHA/Medicare) | 32 |
| Breast screening (AIHW/DOHA) | 34 |
| Bowel screening (AIHW/DOHA/Medicare) | 39 |
| Cervical screening (AIHW/DOHA) | 41 |
| Census data (ABS) | 44 |
| National Health Survey (ABS) | 47 |
| National Aboriginal and Torres Strait Islander Health Survey (ABS) | 51 |
| Survey of Disability, Ageing and Carers 2003 (ABS) | 55 |
| Survey of Mental Health and Wellbeing (ABS) | 58 |
| BEACH survey (University of Sydney/AIHW) | 61 |
| Medicare Service Incentive Payments for cervical screening (DOHA/Medicare) | 63 |
| Palliative Care Outcomes Collaboration data (University of Wollongong) | 64 |
1. Overview

There is a greatly increased demand for national data for all aspects of cancer control, including research, surveillance, service planning and evaluation, quality audit, assessing toxicity and late effects of new treatments, and monitoring consumer views. With these data, improved treatment protocols can be developed, safety and quality monitoring can be strengthened to ensure maintenance of high clinical standards, longer term toxic effects of new treatments can be identified and treatment regimens modified to minimise these effects, and services can be attuned to consumer needs.

A range of data sets are used to respond to data demands, although data under-utilisation is often evident because potential users are unaware of the existence and whereabouts of data. It would be an oversimplification to suggest that this is the only factor limiting data use. Databases are often funded for very specific purposes only, or to provide feedback to their data contributors, such that broader applications are more difficult. In a federated system of government, data clearance may be required through multiple levels of government and through multiple ethics committees, which may further complicate data access.

The document Data Sets for Cancer Control and Research in Australia collates data sets that are of potential value for cancer epidemiology in Australia, together with some more limited and emerging data sets. A description of the data set and its items, and contact details to use when enquiring into data availability and access are included. Some current and possible example applications for research and cancer control are also included.

This monograph has been developed by Professor David Roder on behalf of Cancer Australia. He has been assisted in this work by the National Cancer Data Strategy Advisory Group (NCDSAG). The contributions of the NCDSAG, as well as the Australian Institute of Health and Welfare (AIHW), Medicare Australia, Department of Health and Ageing (DOHA), Australian Bureau of Statistics (ABS), and the Centre for Health Service Development, University of Wollongong, in preparing this report are gratefully acknowledged.

The document Data Sets for Cancer Control and Research in Australia is consistent with the discussion on data availability in A National Cancer Data Strategy for Australia.

The report should be regarded as a living document, with additional data sets and modifications to be added, as the need arises. In particular, it is expected that bio-specimen data sets will be added in due course.

The data sets listed comprise:

1. those commonly used to address cancer questions;
2. additional data sets emerging for this purpose; and
3. others that would make a more occasional contribution, depending on the questions.

Mostly data are available by age, sex, ethnicity, country of birth, place of residence, and type of cancer. Data access would vary with the data collection and would need to be discussed with custodian organisations, together with aspects like data quality and completeness.
2. Introduction and Purpose

The purpose of this monograph is to ensure that potential data users are aware of the cancer data sets that exist. An important pre-requisite for productive use of data is knowledge of what exists and where.

When required data need to be integrated, this may be possible through data linkage, although gaining approvals for this linkage can be challenging. On other occasions, new data need to be collected at source. Knowing what exists and what is emerging facilitates both the productive use of existing data and the planning of additional collection.

An additional purpose of this monograph is to identify data limitations and gaps, such that key initiatives could be identified to increase data capacity. The data limitations identified through this process and some possibilities for improving data availability for health service administration and research are indicated.

This monograph aims to provide an overview of data availability, rather than explore data-dictionary definitions and other details. Such details would need to be obtained directly from data custodians.
3. Data gaps

A. Person descriptors—Indigenous status

Cancer registries obtain data on Indigenous status by derivation from hospital inpatient records, and sometimes, death records. In general the quality and completeness of these data are poor, such that cancer risks and outcomes cannot be monitored effectively for this sector of the population. This also reduces opportunities to use statistics for advocacy.

B. Cancer descriptors

TNM or equivalent stage

A crucial gap is the lack of population coverage with data on TNM (i.e., Tumour, Node, Metastasis) or equivalent stage and patterns of care. Stage is central to evaluating patterns of care and interpreting survival figures. Data on stage also assist evaluation of screening on extent of disease at diagnosis.

Filling this gap is important if inequalities in service delivery and outcomes across the Australian population are to be identified and rectified, and if there is to be effective safety and quality monitoring.

These data are collected in clinical registries, however, these registries cover only a small fraction of all cancers. A more comprehensive picture will be obtained with stage data collection by population-based registries.

Tumour grade

This is also an important prognostic characteristic not collected by most Australian population-based cancer registries.

Treatment

These data are collected in clinical registries; a more comprehensive picture will be obtained with treatment collection for the whole population by population-based registries.
C. OUTCOMES OF CARE

Post-treatment quality of life and function are important outcomes that are difficult to capture on a day-to-day basis through population-based or clinical registration.

D. RISK

While household interviews and telephone surveys are frequently used to collect data on risk factors for cancer and other health conditions in the Australian population, periodic case-control studies could be used to target cancers of special interest.
4. Data sets

4.1: National Cancer Statistics Clearing House

Contact details
Organisation: Australian Institute of Health and Welfare
Name of data set: National Cancer Statistics Clearing House
Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601; Telephone (61)2 6244 1140; Email: cancer@aihw.gov.au

Description
The Clearing House is administered by the Australian Institute of Health and Welfare (AIHW). It contains data on all newly diagnosed invasive cancers since 1982, excepting non-melanoma skin cancers. Limited data are available for in-situ breast, melanoma and some other cancers at state and territory cancer registries, but are not routinely included in the Clearing House.

Clearing House data are provided by state and territory population-based cancer registries, where notification of cancers is legally mandated from pathology laboratories and other notifiers. Coordination of these registries for national data collection purposes occurs through the AIHW and Australasian Association of Cancer Registries (AACR).

Joint AIHW/AACR reports of these data are published annually and are available through the AIHW website. Data are also available through data cubes, accessible through the AIHW website. These data meet routine national requirements for incidence and survival data, with more specific needs requiring customised data extraction.

Normally, access to Clearing House data for customised extractions requires prior approval of individual state/territory registries. The time needed to obtain this approval should be factored into data requests. In addition, where person identified, or potentially person identifiable data are required, ethics committee approval is needed from the AIHW ethics committees, and potentially ethics committees at a state/territory level, increasing the time required.
Data items

Person characteristics:
- Registration number
- State of registration
- Surname
- First name
- Second name
- Third name
- Sex
- Birth date
- Age at diagnosis
- Country of birth
- Race
- Residential SLA, postcode or both (at diagnosis)
- Death date
- Death cause
- Age at death
- SLA at death

Cancer characteristics:
- Cancer registration number
- Diagnosis date
- Basis of diagnosis
- Primary site (topography)
- Histology type
- Histology behaviour code
- ICD-10 cancer code
- Melanoma thickness (mm)
Data limitations

The following items are not included in the Clearing House data set:

- **TNM or equivalent stage**
  
  These data enhance capacity for meaningful patterns of care studies and interpretation of survival figures. Investigation is required to facilitate stage collection by the registries that supply the Clearing House. In particular, means of cost-effectively collecting distant metastasis status at diagnosis is a priority.

- **Tumour grade**
  
  It is a fundamental prognostic indicator and determinant of treatment pathways. It is included in the ICD-O-3 coding manual for use by cancer registries and would be readily available from routine pathology reporting for most solid tumours.

- **Tumour recurrence**
  
  While a clinical characteristic routinely used by clinical researchers to assess disease-free survival, and trends in post-recurrence survival, this information would be difficult to capture at a population level with the data systems used by registries that supply the Clearing House.

- **Cancer risk factors**
  
  Data on tobacco smoking and other risk factors are not normally collected by cancer registries. Such data could be collected in ad hoc surveys for samples of patients and population controls, such that the attribution of these factors to the cancer burden could be reassessed.

- **Other items**
  
  Post-treatment quality of life and function are important outcome measures. They would be difficult to capture routinely by population-based cancer registries, however, and may best be evaluated through follow-up interview surveys of samples of patients.

  Similarly, consumer views on service access, availability, and appropriateness, and satisfaction with services, would better be assessed through sample surveys.
Example applications

- Routine population health surveillance (e.g., monitoring numbers of cancers and cancer deaths, incidence and mortality rates, prevalence and survival by age, sex, country of birth, geographic location, socio-economic status, and diagnostic period).
- Investigations of real and perceived clusters of cancer (often linked to environmental, occupational and related concerns).
- QA activities (e.g., provision of interval cancer data to screening programs).
- Broad service planning, (including cancer projections to indicate changes in staffing and infrastructure needs).
- Providing representative samples of cancers for patterns of care surveys and research (e.g., aetiological, translational (biological to population health), psychosocial, population-health, health-system and consumer research).
- Provision of cancer outcome data for cohort studies of occupational, environmental, behavioural or other risk factors.
4.2: CLINICAL CANCER CORE DATA SET

(Clinical Cancer Core Data Set Specification—National Health Data Dictionary Version 12
Supplement)—Nationally recommended/variably implemented

Contact details

Organisations:

- Cancer Institute NSW, Cancer Registries and Information Division, PO Box 41, Alexandria, NSW 1435, Telephone 61 2 8374 5600, Email information@cancerinstitute.org.au;
- SA Department of Health, Epidemiology Branch, PO Box 287, Rundle Mall, Adelaide, SA 5000, Telephone 61 8 8226 6000;
- WA Department of Health, Information Centre, PO Box 8172, Perth Business Centre, Perth, WA 6849, Telephone 61 8 9222 4222;
- Queensland Health, QCAT, GPO Box 48, Brisbane, Qld 4001, Telephone 61 7 3234 0111;
- Victorian Department of Human Services, Cancer and Palliative Care Unit, 19th Floor, 50 Lonsdale St, Melbourne, Vic 3000, Telephone 61 3 9096 2136

Names of data sets:

- variable

Description

This is a nationally recommended clinical cancer core data set for generic clinical cancer registration, developed under the aegis of the former National Cancer Control Initiative (NCCI) and released in 2004, along with data dictionary definitions. The objective was the collection of enough data to be informative for service monitoring, without imposing an unsustainable burden of data collection.

As with population-based data in the National Cancer Statistics Clearing House, the registration of clinical cancer data is principally a responsibility of states/territories and is often undertaken independently at an individual health agency or clinic level.

These data sets are described because of national initiatives to standardise their content and state/territory initiatives to extend them. There is as yet no national compilation of these data, as described for population-based data. Requests for clinical data need to be sent to state/territory health authorities and service agencies.

While there is a strong interest in clinical cancer registration, in order to collect stage, treatment and other clinical characteristics not routinely collected by existing population-based registries, and major developments are occurring in NSW, Queensland and Victoria, only a very small proportion of cancers in Australia would be covered by clinical registration at this time.
Registry models vary by state/territory. For example:

- WA and SA have hospital-based clinical registries administered by teaching hospitals but linkable to the State population-based registries. The SA clinical registries were introduced in the 1980s and the WA registries in the 1990s. Data sources include population-based registries, hospital-based electronic information systems, case notes and clinician reporting. These registries are mostly restricted to large teaching hospitals.

- NSW has a pilot clinical registry program including five Area Health Services. An emphasis is placed on electronic data extraction from existing inpatient and other administrative databases. There is an interest in NSW in integrating staging, treatment and related clinical data in a re-engineered State population-based cancer registry.

- QLD has a recently introduced web-based public-sector system, accessible across the public sector and linkable to the State population-based registry. This follows earlier development of a separate lung cancer registry covering the public sector. Moves are underway to extend this data collection to the private sector.

- Victoria has provided for the central collection of MDS clinical items in its State population-based registry and electronic means are being developed to transmit these data. An early priority is being given to gynaecological cancers.

- While clinical registries of a good quality are evident in health services in Tasmania, the ACT and Northern Territory, there appears to have been less emphasis on them in these jurisdictions. A Northern Territory workshop has advocated the central collection of clinical items in the Territory’s population-based registry, although this has not yet commenced.

**Data items**

**Person characteristics:**

- Person identifier
- Family name
- Given names(s)
- Medicare card number
- Address
- Sex
- Birth date
- Death date
- Death cause*
**Provider characteristics:**
- Establishment number

**Diagnosis:**
- Primary site (topography)
- Diagnosis date
- Histology type
- Histopathology grade
- Laterality
- Most valid basis of diagnosis
- Performance status at diagnosis*

**Cancer characteristics at diagnosis:**
- Cancer stage—T
- Cancer stage—N
- Cancer stage—M
- Cancer stage—TNM grouping
- Staging basis
- Staging scheme
- Staging scheme edition number
- Tumour size at diagnosis—solid tumours
- Tumour thickness at diagnosis—melanoma
- Oestrogen receptor status
- Progesterone receptor status
- Number of regional lymph nodes examined
- Number of regional lymph nodes positive

**Initial treatment episode:**
- Target site
- Treatment type
- Treatment intention
- Surgical procedure(s)
- Date of surgery
Data Sets for Cancer Control and Research in Australia: A compendium of current available data collections

- Radiotherapy type
- Radiation dose
- Systemic therapy agent name
- Initial treatment—starting date
- Initial treatment—completion date

**Treatment outcomes:**
- Outcome of initial treatment (complete remission/partial remission/etc.)
- Date of first recurrence
- Region of first recurrence (local/regional/distant)

* Excluded from National Data Dictionary. Data on cause of death enables disease-specific survival, which often is preferred to relative survival in clinical referral settings. Performance status is routinely collected in clinical registries due to its prognostic importance.

### Data limitations

The main limitation is the relatively small number of cancers covered by clinical cancer registration. Clinical registration that includes a range of items equivalent to the Clinical Cancer Core Data Set will be difficult to achieve in small sparsely scattered treatment locations.

### Example applications

- Monitoring patterns of care in relation to clinical guidelines.
- Comparing stage-specific survivals with expectations based on the international experience.
- Assessing economic characteristics of alternative treatment pathways.
- Investigating survival effects of differences in diagnostic and treatment practices.
- Describing patterns of care and survivals for high-risk groups, to address inequalities.
- Supporting assessments of intermediate and long term toxicity effects of different chemotherapy and radiotherapy protocols, to protect patient safety.
- Examining operational characteristics of the health system, including patient flows between the public and private sectors, to find opportunities for system improvement.
- Investigating the translation of bench-top research findings into clinical outcomes.
4.3: AIHW NATIONAL MORTALITY DATABASE

Contact details

**Organisation:** Australian Institute of Health and Welfare

**Name of data set:** AIHW National Mortality Database

**Contact:** The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601; Telephone (61)2 6244 1140; Email: cancer@aihw.gov.au

Description

The database covers all deaths registered in Australia since 1965. The information is provided to the AIHW by state/territory death registrars in de-identified form, after coding by the Australian Bureau of Statistics (ABS). The database includes the following items.

Data items

- Record Id
- Month/year of registration
- State of registration
- Registration district
- Registration number
- Sex
- Age at death
- Age group at death
- Birth date
- Place of residence (State and SLA)
- Place of residence (ASGC 9-digit code)
- Place of residence (RRMA)
- Remoteness of residence (5 fields)
- Occupation
- Country of birth
- Duration of residence in Australia
- Indigenous status
Marital status
Age at first marriage
Date of first marriage
Place of first marriage
Number of children

Death details
Certification details
Hospital
Place of death
Post mortem
Drowning flag
Cancer flag
Maternal death flag
Tuberculosis flag
Leukaemia flag
AIDS flag
Asthma flag
Analgesic nephropathy flag
Asbestosis flag
Drug flag
Mortality Id
Place of occurrence
Activity code
Firearms flag
Underlying cause of death (ICD code) (multiple causes available from 1997)
Data limitations

These mortality data do not have person identifiers. Thus they cannot be used as outcome data for cohort studies or be linked to external databases to create de-identified data sets for broader analyses. With appropriate approvals, and strong governance, this may be possible, although the ABS Act may prevent this approach. Internationally, census data are used for data linkage and research (eg, the UK, Scandinavia, and more recently, New Zealand), and similar possibilities should be explored in Australia.

Example applications

- Routine population surveillance of cancer mortality (eg, numbers of deaths and mortality rates by age, sex, country of birth, geographic location, socio-economic status, service access, and calendar year period).
- Ecological studies of effects of cancer services (eg, effects of screening and treatment service access, availability and participation on mortality outcomes).
- Broad service planning (eg, mortality projections to indicate future staffing and infrastructure needs for hospices and other end-of-life services).
- Occupational and environmental research (eg, investigating associations of cause-specific deaths with occupational characteristics and geographic location).
4.4: AIHW NATIONAL DEATH INDEX

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: AIHW National Death Index (NDI)

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601;
Telephone (61)2 6244 1140; Email: cancer@aihw.gov.au

Description

The NDI covers all deaths in Australia from 1980. The data are provided by state/territory death registrars. ABS coded causes of death are added annually. State/territory death registrars are free to impose restrictions on the use of data they provide to the NDI, which the AIHW is obliged through prior agreement to follow. The NDI is used extensively for cancer survival and cohort studies through linkage with cancer-registry incident data and other files.

Data items

- Surname
- Given names (up to three)
- Sex
- Birth date
- Death date
- Death cause (ICD code)
- State of death registration
- Year of registration
- Registration number

Data limitations

The identifiers included in the NDI are those applying at time of death. There is the potential for sub-optimal sensitivity in matching due to name, as may apply, for example, to females upon marriage, or Aboriginal and Torres Strait Islander peoples, although these effects appear in practice to be small. In addition, deaths occurring in other countries may be missed (eg, as may apply when overseas born residents return to their birth country at time of death).
Example applications

- Linked to cancer registries and other clinical databases, routine monitoring of cancer survival and evaluation of effects on cancer survival of screening, treatment services, and health-system change.

- Linked to research data files, follow-up of death data for clinical and population trials and observational studies.

- Linked to demographic, employee, or risk-factor survey data, assessment of death outcomes in cohorts exposed to specified environmental, occupational, and behavioural and other risk factors.
4.5: AIHW NATIONAL HOSPITAL MORBIDITY DATABASE

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: AIHW National Hospital Morbidity Database

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601; Telephone (61)2 6244 1000; Email: info@aihw.gov.au

Description

This database includes summary de-identified data for hospital separations occurring since 1993–94. The data are provided by states/territories. They are episode centred, not person centred (i.e., if a person is admitted more than once in a 12 month period, more than one separation would be entered). Some private hospital separations are excluded from the database (about 3–4%). The potential for methodological differences should be considered if comparing data between states and territories.

Data items

Person characteristics:

- AIHW Record Id
- Person Id
- Birth date
- Age at death
- Residential location (State/Statistical Division & Sub-division/SLA/ARIA/ASGC/RRMA codes/postcode)
- Usual accommodation type
- Country of birth
- Indigenous status
- Employment status
- Contracted patient
- Referral source
- Funding source
- Insurance status
- Patient election status
**Provider characteristics:**
- Establishment identifier
- Hospital location (State/ASGC/RRMA/ARIA codes)
- Hospital-in-the-home

**Predisposing characteristics (where applicable):**
- Activity when injured
- External cause (number, place)

**Admission:**
- Date
- Urgency
- Mode
- Weight (for babies)

**Diagnosis:**
- Principal diagnosis
- Major diagnostic category
- Australian Refined DRGs, versions 4.1, 4.2, 5.0, 5.1
- Additional diagnoses
- Diagnosis number
- Complications and co-morbidity
- Morphology
- Morphology (neoplasm)
- Mental health flag/legal status
**Treatment:**
- Hospital sector
- Care type
- Procedure (block number, code, flag, number)
- Elective surgery (urgency, indicator procedure, surgical specialty, reason for removal, waiting time at removal)
- Same-day Flag
- Intended/actual bed days
- Psychiatric care days
- Leave days

**Outcomes:**
- Further care referral

**Separation:**
- Date
- Mode

**Data limitations**
These morbidity data do not include person identifiers. Thus they cannot be used as outcome data for cohort studies or be linked to external databases to create de-identified data sets to increase data capacity. Opportunities to link these data should be explored. Data are episode centred, such that individuals can contribute on multiple occasions.

**Example applications**
- Routine monitoring of burden of cancer on the hospital system.
- Assessing in-hospital clinical management of cancer by cancer type, socio-demographic descriptors, insurance type, and hospital type, size and location.
- Assessment of effects of geographically defined risk-factor exposures on risk of hospital admission for cancer.
- Data provision for economics research into cancer management.
4.6: Hospital Admissions Minimum Data Set (for 2007–2008)

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: AIHW National Hospital Morbidity Database

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601;
Telephone (61)2 6244 1000; Email: info@aihw.gov.au

Description

This data set follows nationally agreed recommendations for classifying admitted hospital episodes according to patient type, diagnosis and treatment, such that a range of performance indicators can be produced for performance monitoring. Data extractions will be used by the AIHW for the National Hospital Morbidity Database and by the Department of Health and Ageing for the National Admitted Patient Care Data set that it manages. Items in the minimum data set include the following.

Data items

Person characteristics:

- Person identifier
- Area identifier (unique to hospital/agency) of usual residence (ASGC classification)
- Birth date
- Sex
- Country of birth
- Indigenous status
- Mental health legal status

Provider characteristics:

- Australian state/territory identifier (establishment)
- Establishment sector (public/private)
- Region code

Predisposing characteristics (where applicable):

- Activity when injured
- External cause of injury
- Place of occurrence of external cause of injury
Admission:
- Mode of admission
- Urgency of admission (emergency/elective/etc.)
- Admission date
- Source of referral to public psychiatric hospital
- Patient election status (public/private)
- Patient hospital insurance status
- Funding source for patient
- Inter-hospital contracted patient
- Intended length of hospital stay
- Total leave days (from hospital)

Diagnosis:
- Diagnosis related group
- Major diagnostic category
- Principal diagnosis (including ICD cancer codes)
- Additional diagnoses

Treatment:
- Care type (acute/rehabilitation/palliative/etc.)
- Procedure
- Number of days of hospital-in-the-home care
- Number of qualified days for newborns
- Total psychiatric care days

Outcomes:
- Weight in grams (birth weight) for births

Separation:
- Separation date
- Mode of separation (discharge/transfer/death/etc.)
Data limitations

These data do not include person identifiers. Accordingly the data cannot be used as outcome data for cohort studies or be linked to external data to create de-identified data sets to increase data capacity. Opportunities to link these data to increase data capacity should be explored. The data are episode centred, such that individuals can contribute on multiple occasions.

Example applications

- Monitoring quality of hospital care.
- Routine monitoring of burden of cancer on the hospital system.
- Assessing in-hospital clinical management of cancer by cancer type, socio-demographic descriptors, insurance type, and hospital type, size and location.
- Assessing effects of geographically defined risk-factor exposures on risk of hospital admission for cancer.
- Data provision for economics research into cancer management.
4.7: Hospital Case-Mix Protocol

Contact details

Organisation: Department of Health and Ageing, Research and Statistics

Name of data set: Hospital Case-mix Protocol

Contact: Department of Health and Ageing, GPO Box 9848, Canberra ACT 2601; Telephone 1800 020 103; Email: enquiries@health.gov.au

Description

This data set is used by the Department of Health and Ageing. It includes de-identified demographic, clinical and financial (i.e., charge and benefit) data for hospital in-patient episodes covered by private insurance. The following data items are included.

Data items

Person characteristics:
- Person identifier
- Birth date
- Sex
- Residential postcode
- Mental health legal status
- Inter-hospital contracted patient

Provider characteristics:
- Fund identifier
- Hospital provider number
- Hospital contract status
- Hospital type

Referral source:
- Transferring hospital provider number

Charge/benefit paid:
- Insurance cover
- Total days paid
Accommodation charge/benefit
- Theatre charge/benefit
- Labour ward charge/benefit
- Intensive care unit charge/benefit
- Prosthesis items charges/benefits
- Pharmacy charge/benefit
- Bundled charges/benefits
- Other charges/benefits
- Front end deductibles
- Ancillary cover status
- Ancillary charges/benefits
- Medical item charges/benefits/payment type/gap cover
- Hospital-in-the-home care charges/benefits
- Special care nursery charges/benefits
- Coronary care unit charges/benefits

Admission:
- Date
- Time
- Urgency
- Discharge intention
- Assessment only
- Total leave days
- Non-certified days of stay
- Same-day status
- Hospital-in-the-home commencement/completed date/visit days
Data Sets for Cancer Control and Research in Australia: A compendium of current available data collections

Diagnosis:
- DRG code/version
- Principal diagnosis
- Additional diagnosis codes
- Functional impairment codes

Treatment:
- FIM scores at admission/discharge
- ICU days
- Hours of mechanical ventilation
- Care type
- Procedure codes
- Principal MBS item/date
- Miscellaneous service codes
- Minutes of operating theatre time
- Secondary MBS item numbers
- Number of hospital-in-the-home days
- Total psychiatric care days
- ICU hours
- Palliative care status
- Palliative care days
- Unplanned theatre visit
- Special care nursery hours/days
- Coronary care unit hours/days
- Number of qualified days for newborns

Outcomes:
- Weight (infant/neonate/stillborn)
- Re-admission within 28 days
Separation:

- Date
- Mode of separation
- Time of separation

Data limitations

These data do not include person identifiers. Accordingly the data cannot be used as outcome data for cohort studies or be linked to external data to create de-identified data sets to increase data capacity. Opportunities to link these data to increase data capacity should be explored. The data are episode centred, such that individuals can contribute on multiple occasions.

Example applications

- Assessing hospital management of cancer patients by cancer type, co-morbid conditions, socio-demographic descriptors, insurance cover, charges/benefits and hospital type.
- Data provision for economics research into cancer management.
4.8: PRIVATE HOSPITAL DATA BUREAU

Contact details

Organisation: Department of Health and Ageing, Research and Statistics

Name of data set: Hospital Case-mix Protocol

Contact: Department of Health and Ageing, GPO Box 9848, Canberra ACT 2601; Telephone 1800 020 103; Email: enquiries@health.gov.au

Description

This data set includes de-identified data on private hospital separations, including patient demographics, episode data, clinical information, and hospital charges for all patients in private hospitals and day facilities. The collection is administered by the Department of Health and Ageing. The following data items are included.

Data items

Person characteristics:

- Medical record number
- Surname
- Given name
- Sex
- Residential postcode
- Inter-hospital contracted patient
- Mental health legal status

Payment:

- Episode
- Insurance fund membership
- Payer
- Fund/Group identifier

Provider characteristics:

- Hospital type
- Provider number
Referral source:
- Hospital provider number
- Referring facility category

Charge/benefit paid:
- Accommodation charge
- Theatre charge
- Labour ward charge
- Intensive care unit charge
- Prosthesis charge
- Pharmacy charge
- Other charges
- Bundled charges
- Miscellaneous service codes
- Hospital-in-the-home care charges
- Special care nursery charges
- Coronary care unit charges

Admission and stay:
- Date/time
- Urgency
- Discharge intention
- Leave days
- Non-certified days of stay

Diagnosis:
- DRG code/version
- Principal diagnosis code (ICD code/version)
- Additional diagnosis codes
- Palliative care status

Treatment:
- Procedure codes
- Principal MBS item number/date
Data Sets for Cancer Control and Research in Australia: A compendium of current available data collections

- Secondary MBS item number
- ICU hours/days
- Operating theatre time (mins)
- Psychiatric care days
- Care type (e.g., acute/rehabilitation/palliative/geriatric/psychogeriatric/maintenance etc.)
- Unplanned theatre visit
- Hours of mechanical ventilation
- Number of hospital-in-the-home days (including commencement/completion dates)
- Number of hospital-in-the-home visit days
- Number of palliative care days
- Special care nursery days/hours
- Coronary care unit days/hours
- Number of qualified days for newborns

Outcomes:
- Re-admission within 28 days
- Weight (infant/neonate/stillborn)

Separation:
- Date/time
- Mode
- Hospital provider number to which transferred

Data limitations
These morbidity data have the same limitations as those for Data set 7, although including person identification. The data would be episode centred, such that individuals can contribute on multiple occasions.

Example applications
- Assessing hospital management of cancer patients by cancer type, co-morbid conditions, socio-demographic descriptors, insurance cover, charges/benefits and hospital type.
- Data provision for economics research into cancer management.
4.9: HOSPITAL OUTPATIENT CARE MINIMUM DATA SET

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: Hospital Outpatient Care Minimum Data Set

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601;
Telephone (61)2 6244 1000; Email: info@aihw.gov.au

Description

This data set was established in 2005–06 for use in designated outpatient clinics. The data are limited to aggregate counts of occasions of service, although data relating to specified clinic types could have some limited application in cancer research.

Data items

Provider characteristics:

- Australian organisation/establishment identifier
- Outpatient clinic type—allied health, dental, gynaecology, obstetrics, cardiology, endocrinology, oncology (oncology, lymphoedema, radiation oncology), chemotherapy, respiratory, gastroenterology, medical, etc.

Treatment/management:

- Number of occasions of service
- Number of group sessions

Data limitations

These data are limited to aggregate counts and do not include diagnostic codes. They would cover a small and potentially decreasing proportion of outpatient care, since more of this care is being provided in community settings. The role of these data would be limited to providing an outpatient activity indicator.

Example application

- Providing indicative data on extent of cancer management in hospital outpatient settings (eg, for assessing the economic burden of cancer or the impact of policies to increase outpatient as opposed to inpatient care, although it should be noted that cancer care is increasingly being provided in private community settings and that the service contribution of hospital outpatient departments to outpatient care has diminished).
4.10: Medicare Australia Data Sets

Contact details

Organisation: Medicare Australia

Name of data set: Medicare; Pharmaceutical Benefits Scheme

Contact: Client Liaison Unit; Medicare, GPO Box 1001, Tuggeranong, DC, ACT 2901; Telephone 1 800 101 099; Email: statistics@medicareaustralia.gov.au

Description

Medicare Australia administers funding of Australian health services eligible for reimbursement through the Medical Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS). These services exclude those qualifying for benefits through the Department of Veterans’ Affairs, although similar data are collected in the funding of these benefits as for the MBS and PBS. Apart from administering payments, Medicare Australia aims to improve health status through provision of information to health-care decision makers.

There are constraints to information provision, in that release of personal information is subject to secrecy provisions of the National Health Act, the Health Insurance Act and the Privacy Act. In general, de-identified aggregated data can be released where individuals cannot be identified or re-identified. Requests for data are processed through the Client Liaison Unit.

Data items include, but are not limited to, the following.

Data items

Person characteristics:
- Person identifier
- Residential address
- Birth date
- Sex

Provider characteristics:
- Person identifier
- Address
- Specialty by qualification
- Major specialty
- Sex
Medical Benefits Schedule:

- Medicare benefit
- Item number
- Period of service/dates of request/referral
- Whether service provided in hospital

Pharmaceutical Benefits Scheme:

- Medicine (i.e., item code, generic name, brand, strength and quantity)
- Cost
- Original or repeat prescription
- Date of supply
- Payment category
- Number of repeats

Data limitations

Medicare data do not include items on cancer type or co-morbidity. However, as a data source for de-identified data sets that include linked data from cancer registries, hospital inpatients, public radiotherapy centres, and cancer screening registries, they potentially have an important role to play in closing data gaps such that cancer survivals may be analysed by treatment modality. In addition, patterns of care could be investigated to assess appropriateness of care. The inclusion of TNM or equivalent stage on cancer registries (and thus the Cancer Statistics Clearing House) would greatly strengthen the value of these data. Meanwhile, linkage with hospital inpatient records and death records by cause would enable long-term side effects of treatment to be monitored.

Example applications

- Provision of data on radiotherapy, chemotherapy and hormone therapy provided through the private health sector.
- Potentially, through linkage with cancer registries, providing data on surgery, radiotherapy, chemotherapy, hormone therapy and other care delivered through the private sector by cancer type and stage (eg, to assess appropriateness, quality and economic characteristics of care).
- Potentially, through linkages with cancer registries, hospital admission and death record data, providing data on chemotherapy and other treatments for assessing intermediate and long-term toxicity and late effects.
- Potentially, through linkages with cancer registries, providing data for economic studies of alternative models of care.
4.11: BREAST SCREEN DATA

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: BreastScreening Monitoring Data

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601; Telephone (61)2 6244 1140; Email: cancer@aihw.gov.au

Description

The AIHW obtains de-identified unit record data from BreastScreen Screening and Assessment Services for service monitoring and evaluation. Access to these data requires approval from the state/territory screening services providing the data.

Data items (AIHW)

Person characteristics:

- Identification number
- Birth date
- Age at screen
- Postcode of residence
- Indigenous status
- Language spoken at home

Screening characteristics:

- Date of attendance for screen
- Attendance round (1, 2, 3, etc.)
- Recall to assessment status (i.e., not recalled, recalled for mammographic abnormality/other)

Cancer characteristics:

- Tumour size (mm)
- Screen-detected cancer status (none/invasive/DCIS)
Other Data items (these are collected by all or some state/territories)

Identifier:
- Client number
- Screening unit
- Assessment unit
- Surgical unit
- Lesion number
- Service provider
- Machine
- Estimated date flag
- State

Client:
- Name
- Area of residence
- Family history of breast cancer—relationship/age/laterality
- Personal history of breast cancer—year/laterality
- Mammographic history at first screening visit (including year of mammography)
- Round number—state/territory; national program
- Symptom status
- General practitioner flag

Screening visit:
- Booking date
- Date of first attendance (for episode)
- Number of films used
- Technical repeat
- Number of technical repeats
- Screening mammogram reading results
- Screening recommendation
Date woman notified of screening results
Letter to GP about screening results (including date)

**Assessment visit:**
- Reason for assessment
- Attendance
- Date of first attendance for assessment
- Nature of mammographic lesion (including side)
- Nature of signs/symptoms (including side)
- Result of mammography
- Result of clinical examination
- Correspondence relating to clinical examination of mammographic abnormality
- Ultrasound result/description of lesion
- Fine needle cytology method/result
- Core biopsy method/result
- Other assessment procedures
- Final result
- Recommendation (including number of months/date made)
- Assessment visit date
- Discharge from program after assessment
- Date woman notified of assessment results
- Letter to GO about assessment results (including date)

**Local excision of lesion:**
- Local excision performed
- Date of local excision
- Funding
- Marking method
- Localisation technique
- Frozen section
- Specimen x-ray
Lesion identified in specimen
Further surgery recommended
Local excision result
Number of excisions
Date of definitive diagnosis
Recommendation—definitive
Discharge from program after local excision

**Histopathology:**
- Reason for histopathology examination
- Date of diagnosis
- Cancer diagnosed in Program
- Axillary dissection
- Sentinel node biopsy performed
- Axillary dissection—total number of nodes examined/ number positive
- Histopathology of non-malignant/malignant lesions
- Tumour size
- Histological grade
- Dominant lesion identifier

**Primary treatment:**
- Primary treatment form
- Date of treatment commencement
- Side of malignancy
- Surgical treatment
- Adjuvant radiotherapy
- Metastasis—distant (including site)

**Data limitations**

Nationally collected data are de-identified, such that data linkage to cancer registrations, death records or other databases cannot be performed
Example applications

- Assessing service quality (e.g., as required for national accreditation).
- Evaluating effects of screening (e.g., on breast-cancer mortality and prognostic indicators).
- Monitoring screening risks.
- Assessing screening participation.
- Service planning (e.g., assessing future staffing and infrastructure requirements).
- Provision of data for clinical and health-services research
4.12: NATIONAL BOWEL SCREENING REGISTER

Contact details

Organisation: Medicare Australia

Name of data set: National Bowel Screening Register

Contact: Client Liaison Unit; Medicare, GPO Box 1001, Tuggeranong, DC, ACT 2901; Telephone 1800 101 099; Email: statistics@medicareaustralia.gov.au

Description

Data on bowel screening through the National Program are recorded in a Medicare-based Bowel Screening Register. The Register is used for to facilitate screening invitations and follow-up. It shows screening coverage and outcomes of screening and follow-up investigations, such as a positive FOBT results and identification of pre-cancerous polyps, adenomas, and cancers. Data extracts are analysed to monitor and evaluate the Program by the AIHW.

Data items

Person characteristics:

- Participant identifier
- Age at FOBT invitation (indicating screening round)
- Sex
- State & postcode of usual residence
- Indigenous status
- Disability status
- Preferred language for correspondence

FOBT details:

- Screening round
- Sample status (correctly/incorrectly completed)
- Result (positive/negative/inconclusive)

FOBT positive:

- Symptoms at GP assessment
- Colonoscopy referral
- Reason for non-referral (if applicable)
- Referral for other examination
Colonoscopy report:

- Sedation risk
- Sedation used
- Quality of bowel preparation
- Depth of colonoscope insertion
- Visualisation of colon
- Documentation of depth of insertion
- Colonoscopy withdrawal time
- Abnormalities found (including numbers and sites of 2 most advanced cancers, number and size of polyps, and other diagnoses—haemorrhoids, inflammatory bowel disease, diverticular disease, other (specified))
- Adverse outcomes (bleeding, infection, perforation, reaction to sedation, death)

Histopathology report:

- Colorectal polyps (i.e., polyp type, condition, grade)
- Biopsy information
- Local excision or surgical resection of colorectal cancer (i.e., site, histology, grade, extent of invasion, metastatic disease)

Data limitations

Cancer registries have survival data and cancer details, but not screening histories, whereas screening registries lack details of the cancers and survival outcomes. A de-identified linked data set, including cancer registry, hospital inpatient, public radiotherapy, Medicare and screening data, could be used to evaluate the full screening pathway. These data would be strengthened if diagnostic stage were included in cancer registry collections and thereby in the National Cancer Statistics Clearing House.

Example applications

- Monitoring screening participation.
- Assessing service outcomes (eg, removal of pre-cancerous polyps and effects on prognostic characteristics of colorectal cancers).
- Monitoring service quality.
- Potentially, provision of data for evaluating impact of screening on colorectal cancer mortality.
- Potentially, provision of data for service planning and clinical and health-services research
4.13: CERVICAL SCREENING MONITORING DATA
(STANDARDISED CERVICAL SCREENING DATA DICTIONARY)

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: Cervical Screening Monitoring Data

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601;
Telephone (61)2 6244 1140; Email: cancer@aihw.gov.au

Description

Data on cervical screening are recorded by Australian state/territory cervical cytology registries. The data are limited to items recorded on cytology and histopathology reports. The registries are used to invite women for screening when screening has lapsed and to provide safety-net reminders when women with screen-detected abnormalities may not have received follow-up attention. The data show screening coverage and screening outcomes (including detection of pre-invasive abnormalities and cancer). A further application is monitoring of screening quality.

The AIHW does not receive unit-record data from the registries. Tabular data are provided to assess participation rates, extent of early re-screening, low-grade and high-grade abnormality detection, incidence of micro-invasive squamous cell carcinomas, and cancer incidence for squamous, glandular, adeno-squamous and other histology types (also available from the National Cancer Statistics Clearing House). Data items provided to the AIHW are a sub-set of the following items recorded by the state/territory registries.

Data items

Person characteristics:

- Person identifier
- Birth date
- Residential address/postcode
- Hysterectomy status

Screening outcome:

- Date of cytology specimen
- Provider details, including occupation
- Test type
- HPV DNA result
Cervical cytology specimen type:

- Squamous cell analysis and result
- Endocervical cell analysis and result
- Other non-cervical cell analysis and result
- Recommendation

Histopathology:

- Procedure
- Cervical specimen histology
- Endocervical histology
- Endometrial histology
- Vaginal histology
- Vulval histology
- Other histology

Data limitations

Cervical cytology registries do not record Indigenous status, country of birth or measures of ethnicity. This limits the extent of screening monitoring and assessment of cancer risk across the population. Indirect evidence of these and other population descriptors can be obtained through ecological studies in which screening participation and risks of cervical abnormalities are analysed by characteristic of SLA or other geographic unit.

Example applications

- Monitoring screening participation.
- Assessing service quality.
- Evaluating effects of screening on removal of high-grade abnormalities.
- Providing data for evaluating impact of screening, and of changes in policy on management of screen-detected abnormalities, on risk of cervical cancer and cervical cancer mortality.
- Evaluation adherence to screening guidelines.
- Potentially, evaluating the effects of HPV vaccination on risk of pre-invasive abnormalities and cervical cancer.
- Evaluating likely effects of changes in screening frequency and other policy options on risk of cervical cancer.
- Assessing outcomes of recommended screening following HPV testing.
- Provision of data for research on the epidemiology of pre-invasive cervical abnormalities.
- Provision of data for service planning and clinical and health-services research.
4.14: ABS CENSUS

Contact details

Organisation: Australian Bureau of Statistics (ABS), National Information and Referral Service (NIRS)

Name of data set: ABS Census—2006

Contact: The Australian Bureau of Statistics, ABS House, 45 Benjamin Way, Belconnen ACT 2617; Telephone 1300 135070; Email: client.services@abs.gov.au

Description

The ABS undertakes a census of the Australian population five-yearly to obtain population counts and descriptive characteristics. Details are not available at a unit-record level for external users. The data are used by the ABS to indicate the size of the residential population. This information is essential for determining cancer rates. Data items collected in the 2006 census included the following.

Data items

- Place of residence (mesh blocks)
- Place of residence (1yr/5yrs ago)
- Date of birth or age at last birthday
- Sex
- Marital status
- Country of birth
- Year first arrived to live in Australia
- Country of birth of father/mother
- Ancestry
- Language spoken at home
- Proficiency with English
- Citizenship
- Indigenous status
- Religion
- For females, numbers of babies
- School/educational institution attendance
Highest level of school completion
Educational qualification completed/main field of study
Personal income
Employment status/job type/occupation/industry and business type/hours worked
Means of travel to work
Work seeking
Unpaid domestic work
Unpaid care of persons(s) afflicted by disabilities, illnesses and/or old age problems
Unpaid care of child/children
Voluntary work through an organisation or group
Household size and relationships
Number of registered motor vehicles owned/used by residents that were parked at dwelling on census night
Number of bedrooms in dwelling
Dwelling tenure
Rent/other payment for dwelling
Internet access and type at dwelling
Need for assistance with self-care, moving around, communication, etc.
Conditions (reasons) for need of assistance

Data limitations

Numerator information on Indigenous status, ethnicity and other socio-demographic characteristics often are unavailable, or have not been collected in a manner consistent with ABS methodologies, thereby preventing accurate determinations of population rates. This may be countered to some extent through ecological studies at the SLA or other geographic unit level. ABS mesh blocks facilitate the use of small geographic units, which would be of value where cancer data are geo-coded for deriving small area statistics. Direct linking of ABS data to cancer and other external files is prohibited by the ABS Act, although some limited linking might be possible in the 18 months immediately following a census before names are removed. The potential for linking to strengthen the epidemiological data on cancer and cancer risk factors among Aboriginal and Torres Strait Islander peoples and other socio-demographic sub-groups warrants further consideration.
Example applications

- Providing the population data needed to assess cancer rates.
- Enabling ecological research into effects of cancer services on cancer risk and outcomes of cancer (eg, through provision of small-area demographic statistics).
- Potentially, if a mechanism can be found that meets privacy restrictions of the ABS Act, allowing assessments of cancer risk and cancer outcomes for Aboriginal and Torres Strait Islander peoples, and other socio-demographic sub-groups, through linkage with cancer registry data.
4.15: ABS NATIONAL HEALTH SURVEYS

Contact details

Organisation: Australian Bureau of Statistics

Name of data set: ABS National Health Survey

Contact: Health and Disability Section, ABS House, 45 Benjamin Way, Belconnen ACT 2617; Telephone (61)2 6252 6535; Email: client.services@abs.gov.au

Description

National Health Surveys are implemented at three-year intervals through household interview. Their purpose is to collect population data on health status, the prevalence of health-related lifestyle and other risk factors, the use of health services, and frequency of other recent actions undertaken for health reasons. All states and territories and all ages are covered. Notable exclusions are people in hospitals, nursing homes, convalescence homes and hospices. The data are self-reported.

The last survey was in 2004–5 and included approximately 25,900 participants. This would equate with around 500 participants with cancer. This number would be adequate for assessing prevalence for all cancer types combined, and some major types, but there would be sample limitations for assessments for other specific types.

Researchers focusing on individual cancers may be able to augment sample sizes by combining data from more than one NHS survey. Data items available from the 2004–5 NHS survey included the following.

Data items

Person characteristics:

- Age
- Sex
- Marital status (social and registered)
- Country of birth
- Year arriving to live in Australia
- Main language spoken at home
- Proficiency in spoken English
- Relationship in household
- Educational—highest year of school completed, level of highest non-school qualification, type of educational institution attending, main Field of study
- Employment status
Employment details (industry/occupation/etc.)

Personal income—gross cash weekly income, income source, type of government pension/allowance

Tenure, landlord type

Household income

For those aged under 18 years—immunisation status, immunisation details, etc., as reported by adult respondent

Health insurance status, reasons

**Self-assessed health status:**

- Health status now (excellent/very good/good/fair/poor)

**Health risk factors:**

- Height, weight and body mass
- Exercise patterns
- Smoker status
- Dietary behaviours (milk/vegetables/fruit)
- Alcohol consumption
- Breastfeeding
- Adult immunisation status

**Health status indicators:**

- Asthma status/medications/ actions/ etc.
- Cancer
  - Skin-checking behaviour
  - History of cancer
  - Type of cancer (i.e., skin—melanoma, BCC, SCC, other; colorectal; breast; prostate; lung; female reproductive; bladder/kidney; stomach; leukaemia; lymphoma; cancer of unknown site; other (specifying)
  - Age when first diagnosed with breast cancer
  - Current cancer status
- History of hysterectomy/removal of ovaries/HRT use
- Heart and circulatory conditions status/medications
Data Sets for Cancer Control and Research in Australia: A compendium of current available data collections

- Arthritis status/medications/health actions
- Osteoporosis status/medications/health actions
- Diabetes status/medications/health actions
- Psychological distress (K10)
- Mental health medications
- Eyesight/glasses or contact lenses/diabetes related/actions
- Hearing conditions
- Other long-term health conditions (i.e., hay fever, sinusitis, emphysema, bronchitis, anaemia, other allergy, epilepsy, fluid problems, hernias, kidney stones, migraine, psoriasis, gastrointestinal ulcers, thyroid problems, tuberculosis, back problems, other (specifying))
- Work-related conditions, injury-related conditions
- Injuries—types, causes, work-related, location, health actions, time lost from work etc., alcohol/other substance status at time of injury

**Health related actions**
- Days off work/study (in last 2 weeks)
- Days of reduced activity (in last 2 weeks)
- Hospital/day clinic visits/public or private/insurance cover/reasons for non-attendance if needed
- Doctor visits
- Dentist and oral health visits
- Other health professionals visits

**Household characteristics**
- State, Capital city/balance of state, Section of state, ASGC remoteness area, ARIA
- Dwelling type
- Household composition
- Household income
- Household tenure type
Data limitations

It is possible through negotiation to vary survey content to address data gaps and augment sample sizes by jurisdiction to gain greater statistical precision. These surveys therefore offer the advantage of flexibility. However, they rely on self-reporting, restrict data linkage, and have a comparatively small sample size. In the 2004–05 National Health Survey, reason for service utilisation was not collected, so service utilisation by cancer patients cannot be classified as due to cancer or otherwise. In the 2007–08 Survey, visits to general practitioners, specialists and other health professionals have been classified according to whether they were due to cancer.

Example applications

- Assessing the prevalence of cancer and co-morbid conditions.
- Monitoring the frequency of behavioural and other cancer risk factors.
4.16: ABS NATIONAL ABORIGINAL AND TORRES STRAIT ISLANDER HEALTH SURVEY

Contact details

Organisation: Australian Bureau of Statistics, National Information and Referral Service (NIRS)

Name of data set: ABS National Aboriginal and Torres Strait Islander Health Survey

Contact: The Australian Bureau of Statistics, ABS House, 45 Benjamin Way, Belconnen ACT 2617; Telephone 1300 135070; Email: client.services@abs.gov.au

Description

The National Aboriginal and Torres Strait Islander Health Survey (NATSIHS) was conducted in 2004–05, in parallel with the 2004–05 National Health Survey. The survey provided information about the health circumstances of Aboriginal and Torres Strait Islander peoples from remote and non-remote areas across Australia, and about how these circumstances have changed compared with results from Indigenous components of the 1995 and 2001 National Health Surveys (NHS(I)). The survey collected population data on health status measures, health service use and other actions people had recently taken for their health. Included are questions on health related aspects of lifestyle, risk factors, and women's health characteristics. Broad information is available for states and by remoteness area (2001 Australian Standard Geographic Classification, ASGC), and depending on the topic covering most age groups. People in hospitals, nursing homes, convalescence homes and hospices were excluded. The data are self reported, except for the measurement of the height and weight of people in remote areas.

The last survey in 2004–05 included approximately 10,439 Aboriginal and Torres Strait Islander peoples of whom 4,114 were under 15 years of age. The questions were asked by personal interview. Respondents were asked if they had ever been told by a doctor or nurse that they had cancer, and the type of cancer they had. Responses were grouped as: (1) Ever told has cancer, still current and long-term; (2) Ever told has cancer, still current but not long-term; and (3) Ever told has cancer, not current. This is adequate for assessing the prevalence for all cancer types combined and is very limited in cross classifying with other variables.

Data items available from the 2004–05 NATSIHS survey included the following.

Data items

Person characteristics:

- Age
- Sex
- Marital status (Social)
- Main language spoken at home
Indigenous status (Indigenous persons only)
Education—highest year completed, non school qualifications, type of educational institution attending, and main field of study.
Employment status—Industry and Occupation
Personal income—gross weekly income, income source, type of government pension/allowance.

Health risk factors:
- Height, weight and body mass (self-reported and measured for remote only)
- Exercise patterns
- Smoking
- Adult immunisation
- Cultural identification—Stressors
- Dietary behaviours (milk/vegetables/fruit/salt intake)
- Breast feeding and breast feeding history
- Alcohol consumption
- Substance use

Health status indicators:
- Health status now (excellent/very good/good/fair/poor)
- Health status now compared with one year ago
- Asthma status/medication
- Cancer
  - History of cancer—Ever told has condition, still current and long-term, ever told has condition, still current but not long term, or ever told has condition, not current
  - Type of cancer (i.e., skin; colorectal; breast; prostate; lung; female reproductive; bladder/kidney; stomach; leukaemia; lymphoma; cancer of unknown site; other (specifying))
  - Age when first diagnosed with breast cancer
- Heart and circulatory conditions/whether used medication
- Arthritis status/whether used medication
- Osteoporosis status/whether used medication
- Diabetes status/whether used medications/health actions
- Renal disease status/whether had dialysis
- Eyesight/glasses or contact lenses/diabetes related/actions
- Hearing conditions
- Other long-term health conditions (i.e., hay fever, sinusitis, emphysema, bronchitis, anaemia, other allergy, epilepsy, fluid problems, hernias, kidney stones, migraine, psoriasis, gastrointestinal ulcers, thyroid problems, tuberculosis, back problems, other (specifying)
- Social and emotional well-being—mental health medications, psychological distress scale.
- Injury-related conditions—types, causes, work-related, location, health actions, alcohol and other substance status at time of injury

**Health related actions:**
- Days off work/study (in last 2 weeks)
- Days of reduced activity (in last 2 weeks)
- Hospital/day clinic visits/public or private/insurance cover/reasons for non-attendance if needed
- Doctor visits/payment/reasons for non-attendance if needed
- Dentist and oral health visits/dental health status/reasons for non-attendance if needed
- Other health professionals visits/reasons for non-attendance if needed
- Discrimination
- Health Insurance

**Women’s and men’s health:**
- History of mammography, Pap smear, reproductive history, use of contraceptives
- Men’s health—consultations

**Household characteristics:**
- ASGC Remoteness area and States
- Dwelling structure
- Household composition
- Household income
- Household tenure type
Data limitations

As with National Health Surveys in general, it is possible through negotiation to vary survey content to address data gaps. Again, the surveys include reliance on self-reporting, restrictions on data linkage to check accuracy of self-reporting, and very small sample sizes of Aboriginal and Torres Strait Islander peoples for assessing low-incidence conditions, such as most site-specific cancers. Also, reasons for service utilisation were not collected in the 2004–05 Survey, such that services cannot be classified as due to cancer or otherwise.

Example applications

- Assessing the prevalence of cancer and co-morbid conditions.
- Monitoring the frequency of behavioural and other cancer risk factors.
- Evaluating screening and other service utilisation by people affected by cancer.
4.17: ABS SURVEY OF DISABILITY, AGEING AND CARERS

Contact details

Organisation: Australian Bureau of Statistics

Name of data set: ABS Survey of Disability, Ageing and Carers

Contact: Health and Disability Section, ABS House, 45 Benjamin Way, Belconnen ACT 2617; Telephone (61)2 6252 6535; Email: client.services@abs.gov.au

Description

Cancer data were collected in disability, ageing and carer surveys in 1998 and 2003. In the 2003 survey, the principal aims were to determine the prevalence of disability, the need for support of older people and those with a disability, and the characteristics of carers. Long-term health conditions (i.e. conditions lasting or likely to last for 6 months or more, or causing an impairment of this duration) were reported.

Data collection was through a combined household survey and a cared-accommodation survey. The cancer data collected in 1998 were categorised in available unit-record files as cancers of the skin, breast, or other types, whereas those collected in 2003 were classified as breast cancer, prostate cancer or other neoplasm. As in other interview surveys, the data collected on cancer are limited by sample size.

Data items

Person characteristics:

- Age
- Sex
- Marital status
- Number and age of children
- Age of youngest dependent child
- Employment status, occupation, industry
- Country of birth
- Education
- Housing tenure
- Income
- Transport used
Household characteristics:
- Dwelling information
- Household type
- Number of people by age in household, family relationships, number with a disability or who is a carer
- Family and household income
- State
- Indices of disadvantage

Self-assessed long-term conditions (incl Cancer):
- Age at onset
- Time from onset
- Cause of main condition
- Whether completed non-school qualification before onset of main condition
- Condition type (e.g., cancers of the colon, lung, skin, breast, prostate, brain, other malignant/ benign, Hodgkin’s disease, lymphoma, leukaemia)

Self-assessed disability:
- Status—Core activity limitations with communication, mobility and self care and restrictions with education or employment
- Type—sensory/speech, intellectual, physical, psychological, head injury, stroke or brain damage
- Extent/effects—education, employment, labour force participation
- Main condition associated with disability
- Aids—communication, home modifications, medical mobility, self-care
- Need for assistance—communication, cognition/emotion, health care, housework, meal preparation, mobility, paperwork, property maintenance, self care, transport
- Assistance received—cognition/emotion, communication, health care, housework, meal preparation, mobility, paperwork, property maintenance, self-care, transport
- Care provider
- Care status
- Community participation—activities at home, activities away from home, computer use
Carers:
- Assistance provided, status, family level, household level, income level, recipient information
- Effect of caring role on carers physical, financial and emotional wellbeing and access to support and assistance

Data limitations
The limitations of these surveys include uncertainties of self-reporting, restrictions on data linkage to check accuracy of self-reporting, and small sample sizes that would limit analyses by cancer type. Breast cancer was collected in both the 1998 and 2003 survey and it may be possible to combine items from both surveys to increase sample size, although this could be a difficult process.

Example applications
- Assessing statistical associations of cancer with disability.
- Evaluating support needs of people with cancer-related disabilities and carer availability, care provision, and carer characteristics, although the utility of analyses may be restricted by small numbers.
### 4.18: ABS Survey of Mental Health and Wellbeing

**Contact details**

**Organisation:** Australian Bureau of Statistics  
**Name of data set:** ABS Survey of Mental Health and Wellbeing  
**Contact:** Health and Disability Section, ABS House, 45 Benjamin Way, Belconnen ACT 2617; Telephone (61)2 6252 6535; Email: client.services@abs.gov.au

**Description**

The Survey of Mental Health and Wellbeing was conducted in 1997 and 2007. The 2007 survey data has been released. The main purpose of the survey is to measure the prevalence of mental health conditions in Australia, but it also provides information on co-morbidities (including chronic conditions, such as cancer), functioning, and service utilisation. The survey is based on the World Mental Health Composite International Diagnostic Interview (WMH-CIDI). The content of the WMH-CIDI has been modified in part for Australian conditions.

The data are self-reported. The 1997 survey included 10,641 participants. This number would be adequate for analysing the comorbidity of having cancer and a mental health condition, but there are sample limitations in examining the comorbidity of cancer and specific mental health conditions. It is anticipated that this will also be the case for the 2007 survey.

The surveys cover all states and territories but exclude very remote areas. Other notable exclusions are people living in non-private dwellings (such as hostels, boarding houses and institutions), people from overseas holidaying in Australia, members of non-Australian defence forces and their dependents stationed in Australia, households containing non-Australian diplomatic personnel and homeless people.

**Data Items**

**Cancer specific**

In the 1997 survey, participants were asked about their physical health, and were asked to answer "yes" or "no" to whether they currently had any medical conditions that usually last for some time. If they responded "yes", they were asked to provide details, and cancer was one of the possible responses. The information reported was not medically verified, and was not necessarily based on diagnoses by a medical practitioner.

In the 2007 survey, participants were asked whether they had ever been told by a doctor or nurse that they had/have any type of cancer. For participants who responded "yes", they were also asked if the cancer ever lasted for six months or more, and how old they were the first time they had cancer.
Person characteristics:

- Age
- Sex
- Marital status (social and registered)
- Country of birth
- Year of arrival
- Main language spoken at home
- Proficiency in spoken English
- Relationship in household
- Father’s country of birth
- Mother’s country of birth
- Time since end of last relationship

Education—Whether currently studying full-time or part-time, Level of highest educational attainment, Main field of highest educational attainment, Highest year of school completed

Labour force—Labour force status, Occupation (main job), Hours usually worked each week (all jobs), Reason for not being in the labour force

Personal income—Personal gross weekly cash income, Source(s) of personal cash income, Main source of personal cash income, Receipt of government pensions/allowance

Perception of prosperity
- Sexual orientation
- Veteran’s indicator
- Long-term health conditions
- Co-morbidity
- Disability
- Exercise status
- Smoker status
- Body Mass Index
- Substance use
- Medication use
- Symptoms of use
Suicidality
Service utilisation
Social networks
Caregiving

Household characteristics:
- Family composition
- Number of persons in household
- Housing tenure type
- Equivalised income
- Household income
- Sources of household cash income
- Financial strain/stress
- Index of economic resources
- Index of disadvantage

Data limitations
The limitations of these survey data include reliance on self-reporting, restrictions on data linkage to check accuracy of self-reporting, limited data provision on cancer status, and small sample sizes.

Example applications
- Assessing statistical associations of cancer with mental health.
4.19: AIHW CONTINUOUS NATIONAL SURVEY OF GENERAL PRACTICE ACTIVITY (BEACH)

Contact details

Organisation: Australian Institute of Health and Welfare

Name of data set: BEACH survey

Contact: The Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601; Telephone (61)2 6244 1000; Email: info@aihw.gov.au

Description

An ongoing sample survey of general practice activity is used to collect data on patient risk factors and health status. The data are owned by the General Practice Statistics and Classification Unit (University of Sydney), which will consider data requests. The survey is part of the BEACH study (Bettering the Evaluation and Care of Health).

Data items

GP characteristics:
- Age
- Sex
- Years in general practice
- Sessions per week
- Country of graduation
- Practice size
- Computer use
- Hours worked/on call per week
- Location (i.e., rurality) of practice

Patient characteristics:
- Age
- Sex
- Non-English Speaking Background status
- Aboriginality
- Torres Strait Islander status
Health Care Card and Veteran Affairs status
Status to the practice (new/seen before)

**Encounter:**
- Characteristics
- Reasons
- Problems managed
- Type of management
- Tests and investigations
- Risk factors (smoking, alcohol intake, body mass index, etc.)

**Data limitations**
These data provide a “window” on cancer risk behaviour among patients attending general practices covered by the survey. The sample sizes would be limited for detailed applications and the results should be regarded as indicative rather than comprehensive and definitive.

**Example application**
- Assessing the prevalence of cancer risk behaviour by patient characteristic.
4.20: MEDICARE SERVICE INCENTIVE PAYMENTS FOR CERVICAL SCREENING

Contact details

**Organisation:** Department of Health and Ageing, Research and Statistics

**Name of data set:** Medicare Service Incentive Payments for Cervical Screening

**Contact:** Department of Health and Ageing, GPO Box 9848, Canberra ACT 2601; Telephone 1800 020 103; Email: enquiries@health.gov.au

Description

Medicare Australia provides data on Service Incentive Payments (SIPs) for cervical screening by state, Division of General Practice, and Rural, Remote and Metropolitan Areas (RRMA) classification categories, which could have a research application. The data relate to screens for women who have not been screened for four years or more and are screened by a practitioner registered with the PIP program.

Data items (as provided in tabular reports)

- Number of PIP practices “signed on” for PIP
- Number “signed on” for cervical screening
- Number of eligible providers
- Number of SIP services
- % of “signed on practices” with SIPs
- % of eligible providers receiving SIPs

The relevant GP item numbers in relation to cervical screening are 2497, 2501, 2503, 2504, 2506, 2507, 2509, 2598, 2600, 2603, 2606, 2610, 2613 and 2616.

Data limitations

These are summary tabular data only, which limits their utility to assessing broad service activity.

Example application

- Assessing affects of screening promotion on screening activity for high-risk women
4.21: Palliative Care Outcomes Collaboration

Contact details

Organisation: Palliative Care Outcomes Collaboration (PCOC)

Name of data set: Palliative Care Outcomes Collaboration (PCOC)

Contact: PCOC Manager, Centre for Health Service Development, University of Wollongong, NSW 2522; Telephone (61)2 4221 4411

Description

The PCOC has been funded by the Australian Government from 2003–2008 Health Care Agreements to improve palliative care practice and meet “Standards for Providing Quality Palliative Care for all Australians”. A key element is data collection and national benchmarking. All palliative service providers in Australia are eligible to participate on a voluntary basis. The four principal funded partners are the University of Wollongong, Curtin University of Technology, Flinders University, and the Queensland University of Technology. Data items collected include the following.

Data items

Patient characteristics:

- Unique identifier
- Birth date
- Sex
- State/Territory and postcode of residence
- Indigenous status
- Main language spoken at home
- Country of birth

Episode characteristics:

- Referral date
- Referral source
- Commencement date (first contact/care commencement)
- Accommodation type at episode start
- Accommodation type at episode start
- Mode of start
Type
- Consultation and/or assessment only
- Location of consultative service
- Mode of consultative service (i.e., face-to-face, telephone/electronic communication)
- Diagnosis (i.e., malignancy type, other)
- Accommodation type at episode end
- Accommodation type at episode end
- Mode of end (stating place of death, if applicable)
- End date

Phase characteristics:
- Date of start
- Phase type (i.e., stable, unstable, deteriorating, terminal, bereavement support))
- Model type (i.e., direct care, shared care with GP, shared care with other provider (i.e., cancer, respiratory etc.))
- Functional scores at phase start (i.e., RUG-ADL scores)
- Symptom Assessment Score at phase start
- Problem severity scores at phase start (i.e., for pain, other symptoms, psychological/spiritual, family/carer, etc.)
- Karnofsky functional score at phase start
- Provider type
- Date of end
- Reason for end (i.e., change in condition, discharge/case closure, death, bereavement phase end)
- Functional scores at phase end (i.e., RUG-ADL scores)
- Symptom Assessment Score at phase end
- Problem severity scores at phase end (i.e., for pain, other symptoms, psychological/spiritual, family/carer, etc.)
- Karnofsky functional score at phase end
- Number of days seen
- Total leave days
Data limitations

These data relate to PCOC participants only. It is uncertain whether they would be representative of palliative care services provided by non-participants.

Example application

- Assessing episode and phase characteristics of palliative care by socio-demographic characteristic, cancer type, etc.