

# NSW Central Cancer Registry Data Dictionary

## Cancer Institute NSW

### Background

The Central Cancer Registry (CCR) receives notifications of cancer in NSW. It is managed by the Cancer Institute NSW for the NSW Department of Health (DOH), and operates under the authority of the *Public Health Act 1991*. The Registry maintains a record of all cases of cancer diagnosed in NSW residents since 1972 and a record of all cases where a person has died from cancer from 1972 onwards.

The CCR is a case-based registry in which all notifications relating to a particular cancer are linked to a single person. A cancer is the occurrence of a primary malignant neoplasm in an organ of a particular person. If the same person has another cancer, that cancer counts as a second case. The CCR records new cancer cases but not cancer recurrences. Deaths are recorded for all people dying from cancer whose cancer was diagnosed when they were residing in NSW, even if the death occurs outside NSW.

### Coverage

The NSW CCR was established in 1971 as a population-based register of all cancers in NSW residents. The data relate to invasive primary cancers diagnosed in NSW since 1972, and cancer deaths. It does not include skin cancers other than melanoma (e.g. basal cell carcinoma and squamous cell carcinoma are excluded).

Notification of malignant neoplasms and pre-malignant breast carcinoma and melanoma is a statutory requirement for public and private hospital, departments of radiation oncology, nursing homes, pathology laboratories, outpatient departments and day procedure centres. Mandatory pathology reporting has occurred since 1986. Data collected by the Central Cancer Registry includes personal identifiers and demographic information, brief clinical details describing cancer and a record of at least one episode of care from a notifier. Similar details are supplied electronically by many hospitals. The data are supplemented by information from pathology reports and death certificates. In situ lesions for breast cancer and cutaneous melanoma are also routinely collected.

### Quality

Routine data quality control measures conducted include: monitoring notification rates by notifier to ensure that notification activity is being maintained, data entry validation checks and cross checks of data items, periodic checks of coding accuracy and reliability, reconciliation of information from multiple sources, thorough examination of multiple registrations, collaboration with medical experts and other cancer registries, International Agency for Research in Cancer program data check and use of IACR conventions.

### Using Central Cancer Registry data

A final dataset is produced each year as soon as data collection for all cases and deaths for that year are finalised and processed. This final dataset is completely validated and cross checked and is uploaded for general internal utilization. This dataset is also used to produce the annual Cancer Incidence and Mortality report, Central Cancer Registry web based reporting module and the web cube. Any interim datasets produced will not be validated or cross checked.

For data access enquiries please Email: [DARenquiries@cancerinstitute.org.au](mailto:DARenquiries@cancerinstitute.org.au)

## Timeline of CCR milestones

Year	CCR Milestone
1971	Dr Joyce Ford appointed with state funding to establish the NSW Central Cancer Registry in the NSW Department of Health
1972	Registry Commenced
1976	First full report on cancer and incidence mortality for 1972 published
1986	Registry transferred to the administration of The Cancer Council NSW
1986	Pathology notifications commenced
1990	The Cancer Epidemiology Research unit formed
1990	Notification by pathology laboratories became mandatory under the Public Health Act
Mid 1990s	Registry moved from manual to partial electronic notification of inpatient notifications and linked the Registry's database to the NSW and National Death Indexes
2000	ICDO-2 commenced
2000	In situ breast collection commenced
2001	Registry introduced a Work Flow Management System and introduced a Reporting module for easier analysis of data
2002	ICDO-3 commenced
2002	Myelodysplasia collection commenced
2004	Registry transferred to the administration of the Cancer Institute NSW
2007	Cancer Notification Portal developed

## NSW Central Cancer Registry - Data Dictionary

ID	Variable	Description/Notes	Data Values
<b>Demographic Data</b>			
2	Sex	Gender of the person	1 - Male 2 - Female
3	Country of Birth	<p>Country of birth – the Standard Australian Classification of Countries (SACC 2<sup>nd</sup> edition). Issued by the Australian Bureau of Statistics. The two digit Country of birth code was replaced in June 2008, two digit codes were translated to SACC codes.</p> <p>This data item is not available on pathology reports notified to the Central Cancer Registry. Particular cancer sites like melanoma where a person does not have a hospital visit and the pathology report is the main source of notification will have a large proportion that are unknown country of birth.</p>	<a href="http://www.abs.gov.au/ausstats/abs@.nsf/mf/1269.0">http://www.abs.gov.au/ausstats/abs@.nsf/mf/1269.0</a>
4	Indigenous Status	<p>Aboriginal and Torres Strait Islander status. Recommended to be used for incidence (with imputation) from 1999 and mortality from 1994.</p> <p>Approximately 20% of cases do not have an aboriginal status recorded on the NSW CCR. This affects primarily incidence calculations. We strongly advise imputation is used when calculating Aboriginal status from NSW CCR data. Details of methods used by the NSW CCR can be found in the report: .....</p> <p>Periods of usage: From 1994 for mortality and 1999 (with imputation) for incidence. This data item has been subject to quality review – please see the link to the following reports: NSW technical report on completeness of the Aboriginal identifier in the NSW Central Cancer Registry</p> <p>National Report on data quality and completeness. Permission is also required from the Aboriginal Health and Medical Research Council when reporting this data item</p>	1 – Aboriginal 2 - Torres Strait Islander 3 - Both 4 –Neither 8 –Declined 9 -Unknown
5	Year of Birth	<p>Year of birth</p> <p>In cases where the age AND Date of Birth were NOT known, the DOB was entered as 17 Nov 1858.</p>	YYYY
6	Month of Birth	<p>Month of birth</p> <p>In cases where the age AND Date of Birth were NOT known, the DOB was entered as 17 Nov 1858.</p>	MM
7	Day of Birth	<p>Day of birth</p> <p>In cases where the age AND Date of Birth were NOT known, the DOB was entered as 17 Nov 1858.</p>	DD
8	Date of Birth Estimated	Flag identifying if date of birth is estimated / calculated from an entered age. Date of Birth was not a mandatory item prior to 1994	Y - estimated
<b>Cancer Diagnosis Data</b>			
9	Year of Diagnosis	<p>Year of diagnosis from the first definitive treatment as determined using the International Association of Cancer Registries definitions. If a pathology report has been received then the specimen date is used as the date of diagnosis.</p> <p>If Year of Diagnosis =1888 then year of diagnosis is unknown</p>	YYYY

ID	Variable	Description/Notes	Data Values
10	Month of Diagnosis	Month of diagnosis as determined using the International Association of Cancer Registries definitions. If a pathology report has been received then the specimen date is used as the date of diagnosis.	MM
11	Date of Diagnosis Estimated	Flag identifying if date of diagnosis is estimated	Y - estimated
12	Age at Diagnosis	Is a calculated field that is the age in years that a person was when they were diagnosed for their cancer case. We do not receive day of diagnosis, therefore it is artificially set to 15.	
13	Cancer Group	The type of cancer according to the reporting categories for cancer. Tabulation category for cancer site derived from ICD-O-3 site and morphology. This categorization was originally developed to allow consistent reporting of cancer registry data – it is based on ICD-9 groupings	See Attachment 1 – Reporting Categories for Cancers
14	ICD-0-3 Topography Code	The anatomical site of the cancer. Coded according to the WHO International Classification of Diseases for Oncology (third edition)	<a href="http://www.who.int/classifications/icd/adaptations/oncology/en/">http://www.who.int/classifications/icd/adaptations/oncology/en/</a>
15	ICD-10-AM Topography Code	A combination of anatomical site and cell type of cancer. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM). Australian modification of the WHO ICD-10 base classification system, modified to serve particular Australian needs and to support the national collection of data relevant to the population's health.	
16	ICD-0-3 Morphology Code (3 Digit)	Morphology codes describe the cell type of cancer according to the WHO International Classification of Diseases for Oncology (third edition). A three morphology code refers to the main groupings by cell type e.g. Lymphoid leukemia's The NSW Central Cancer Registry requires that all invasive neoplasms or cancers should be notified including insitu breast and melanoma cases. Morphology codes have changed over time, more recent changes include an expansion of lymphohaematopoietic malignancies. Myelodysplasia was previously considered a preleukemic condition, however, since 2004 this item has been considered a malignant condition and is therefore notified to the registry and available to researchers from this date	<a href="http://www.who.int/classifications/icd/adaptations/oncology/en/">http://www.who.int/classifications/icd/adaptations/oncology/en/</a>
17	ICD-0-3 Morphology Code (4 Digit)	The description is the same as 16 with the exception that a four digit morphology code contains the more specific code e.g. Acute leukaemia, Not otherwise specified.	<a href="http://www.who.int/classifications/icd/adaptations/oncology/en/">http://www.who.int/classifications/icd/adaptations/oncology/en/</a>
18	Behavior Code	The behavior of the tumor at diagnosis for all invasive cancers and. In-situ breast and melanoma only.	2 – In-situ 3 – Invasive
19	Best Basis of Diagnosis	Method of diagnosis is a summarised field based on the highest level of verification of cancer specimens within a four month period. For example if a case is originally notified through cytology and a pathology report is received and read by a Central Cancer Registry coder, then the method at diagnosis is 6.	0 – Cytology including FNA, smears, washing, sputum 1 – Clinical/imaging/biochemical 2 – Histopathology performed 3 – Confirmed by post mortem (obsolete code) 4 – Found at post mortem 5 – Death certificate only 6 – Histopathology sighted at CCR

ID	Variable	Description/Notes	Data Values
20	Degree of Spread at Diagnosis	Assigned from the highest degree of spread (extent of spread of cancer from its point of origin) within four months of the cancer diagnosis. Refer to caveat document for further detail.	0 – In-situ 1 – Localised to tissue of origin 2 – Regional spread, adjacent organs and/or regional lymph nodes 3 – Distant metastases 9 – Unknown
21	Laterality	Laterality describes which side of a paired organ is the origin of the primary cancer. Each side of a paired organ is considered separately and described as lateral when occurring unless a physician determines that it is bilateral. A paired organ is one in which there are two separate organs of the same kind, one on either side of the body (e.g. kidney, breast, ovary, testis and lung).	L (left) or R (right) B (bilateral)
22	Thickness of Melanomas	Breslow thickness in mm Cutaneous Melanoma thickness is recorded completely for all cases diagnosed in 1988 onwards. Thickness (mm) is the thickness of the total tissue extracted (multis not included).	Valid values are 0.1-994 (mm). Other values are: 995=microinvasion 996=size < 0.1 mm 997 = size described as small 998 = size described as big 999= unknown
23	Size of Breast Cancer	Size in mm For breast, size is recorded completely for all cases diagnosed in 1997 only and from 2004 onwards. Size (mm) is the size of the total breast cancer tissue extracted (multis not included). There are guidelines for recording breast cancer tumour size – please see attached.	Valid values are 0.1-994 (mm). Other values are: 995=microinvasion 996=size < 0.1 mm 997 = size described as small 998 = size described as big 999= unknown
24	Number of Primary Sites	Primary sites are defined as the number of invasive cancer cases reported for statistical purposes in the register. The number of primary sites is subject to the International Association of Cancer Registries Multiple primary rules, which allows for multiple primary sites per person.	
<b>Cancer Death Data</b>			
25	Year of Death	Year of death	YYYY
26	Month of Death	Month of death	MM
27	Day of Death	Day of death	DD
28	Age at Death	How old the person was when they died.	
29	Reporting cause of death grouping	The type of cancer according to the reporting categories for cancer Tabulation category for cancer site derived from ICD-O-3 site and morphology. Use for reporting purposes. See explanation for cancer group, item 13.	See Attachment 1 – Reporting Categories for Cancers.
30	ICD-0-3 Topography Cause of Death Code	The anatomical site of the cancer. Coded according to the WHO International Classification of Diseases for Oncology (third edition). See explanation for ICDO-3 topography code, item 14.	<a href="http://www.who.int/classifications/icd/adaptations/oncology/en/">http://www.who.int/classifications/icd/adaptations/oncology/en/</a>
31	ICD-10-AM Cause of death Code	A combination of anatomical site and morphology code of a cancer. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM). Australian modification of the WHO ICD-10 base classification system, modified to serve particular Australian needs and to support the national collection of data relevant to the population's health. See explanation for ICD-10-AM Topography Code, item 15.	

ID	Variable	Description/Notes	Data Values
32	Place of death	Caveats for this variable are currently under development. The data dictionary will be updated as soon as this information is available. In the meantime if you are interested in obtaining this variable please state this variable in your data request form to us. Following receipt of the data request form we will contact you for further discussion.	
Residential Data at Diagnosis			
32	Postcode	Postcode of residence at diagnosis	
33	LGA	Local Government Area of residence at diagnosis according to the Australian Standard Geographical Classification (ASGC 2006) issued by the Australian Bureau of Statistics. LGA is a geographical area under the responsibility of an incorporated local government council, or an incorporated Indigenous government council.	<a href="http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/BD1B52D132D130E7CA2573630012F67B?opendocument">http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/BD1B52D132D130E7CA2573630012F67B?opendocument</a>
34	SLA	Statistical Local Area of residence <b>at diagnosis</b> according to the Australian Standard Geographical Classification (ASGC 2006) issued by the Australian Bureau of Statistics. SLAs are Local Government Areas (LGAs), or parts thereof. Where there is no incorporated body of local government, SLAs are defined to cover the unincorporated areas.	<a href="http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/BD1B52D132D130E7CA2573630012F67B?opendocument">http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/BD1B52D132D130E7CA2573630012F67B?opendocument</a>
35	ARIA	Accessibility/Remoteness Index of Australia (ARIA) according to the Australian Standard Geographical Classification (ASGC) 2006 Remoteness Structure. Categories range from 'Major Cities of Australia' to 'Very Remote Australia'. ARIA is assigned to records pertaining to <b>year of diagnosis 2000 and onwards</b> .	<a href="http://www.abs.gov.au/website/dbs/D3310114.nsf/home/remoteness+structure">http://www.abs.gov.au/website/dbs/D3310114.nsf/home/remoteness+structure</a>
36	SEIFA	Socio-Economic Indexes for Areas (SEIFA) is a product developed especially for those interested in the assessment of the welfare of Australian communities. The ABS has developed four indexes to allow ranking of regions/areas, providing a method of determining the level of social and economic well-being in each region. SEIFA is assigned to records pertaining to <b>year of diagnosis 1980 and onwards</b> .	<a href="http://www.abs.gov.au/website/dbs/D3310114.nsf/home/Seifa_entry_page">http://www.abs.gov.au/website/dbs/D3310114.nsf/home/Seifa_entry_page</a>
37	Local Health District	Local Health District of residence at diagnosis. NSW Health division of the state public services, as at July 2011. Eight Local Health Districts cover the Sydney metropolitan region and seven cover rural and regional NSW.	<a href="http://www.health.nsw.gov.au/services/index.asp">http://www.health.nsw.gov.au/services/index.asp</a>
38	Area Health Service	Area Health Service of residence at diagnosis NSW Health division of the state public services, as at 2005. Eight Local Health Districts cover the Sydney metropolitan region and seven cover rural and regional NSW.	

## Attachment 1 - Reporting Categories for Cancer Type

<i>Code</i>	<i>Description</i>		
C00	Lip	M95	Other lymphatic haematopoietic
C012	Tongue	0006	Non Cancer Death
C036	Mouth	0007	Unknown cause of death
C078	Salivary glands		
C0910	Oropharynx		
C11	Nasopharynx		
C123	Hypopharynx		
C14	Other oral cavity & pharynx		
C15	Oesophagus		
C16	Stomach		
C17	Small intestine		
C18	Colon		
C1921	Rectum, rectosigmoid, anus		
C22	Liver		
C234	Gallbladder		
C25	Pancreas		
C301	Nose, sinuses, etc		
C32	Larynx		
C334	Lung, bronchus		
C378	Other thoracic organs		
C401	Bone		
C43	Melanoma of skin		
C44	Other skin		
C45	Mesothelioma		
C46	Kaposi's sarcoma		
C479	Connective tissue, peripheral nerves		
C50	Breast		
C53	Cervix		
C545	Uterus, Body & NOS		
C567	Ovary		
C58	Placenta		
C59	Other female genital organs		
C61	Prostate		
C62	Testis		
C6360	Other male genital organs		
C648	Kidney, etc		
C67	Bladder		
C69	Eye		
C71	Brain		
C72	Central nervous system		
C73	Thyroid		
C745	Other endocrine glands		
C81	Hodgkin's disease		
C82	Non-Hodgkin's lymphoma		
C8890	Multiple myeloma		
C910	Acute lymphoblastic leukaemia		
C919	Other lymphoid leukaemias		
C920	Acute myeloid leukaemia		
C929	Other myeloid leukaemia		
C94	Other specified leukaemias		
C95	Unspecified leukaemias		
C98	Il-defined & unspecified site		