

NSW Cancer Registry – Data Dictionary

Version: 1972-2015 cancer incidence and mortality and 2013 clinical episodes of care from the NSW Cancer Registry (NSWCR)

Item No.	Element	Description/notes	Data values/format
Demographic data elements			
1.	Gender	The gender of the person.	M – Male F – Female
2.	Country of birth	<p>The country in which the person was born as represented by a code.</p> <p>The Standard Australian Classification of Countries (SACC 2nd edition), issued by the Australian Bureau of Statistics.</p> <p><u>Caveat:</u> Country of birth is incomplete on the NSWCR. Country of birth is updated from hospital inpatient records if it is available. Particular cancer sites like melanoma, where a person does not always have a hospital visit and the pathology report is the main source of notification, will have a larger proportion of unknown country of birth. For the most recent 5-year period of data available (2011-2015), approximately 15% of diagnoses had unknown country of birth.</p> <p>Country of birth for persons diagnosed in 2014 were impacted by a system issue, producing an increased proportion of unknown country of birth. The system issue is planned to be rectified for the next cancer incidence and mortality dataset release.</p>	<p>http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/2547A73B811A360DCA25744B00153861?opendocument</p> <p>Unknown COB code = 0000</p>
3.	Aboriginal and Torres Strait Islander status	<p>The Aboriginal and Torres Strait Islander status of the person.</p> <p>Permission is required from the Aboriginal Health and Medical Research Council when publishing reports utilising this data element.</p> <p><u>Caveat:</u> Aboriginal and Torres Strait Islander status is incomplete on the NSW Cancer Registry (NSWCR). This data element is updated from hospital inpatient records if available and ABS coded cause of death data. This affects primarily incidence calculations.</p> <p>From 1999 to 2015, approximately 11% of all NSW invasive cancer cases recorded on the NSWCR, did not have an Aboriginal and Torres Strait Islander status recorded.</p> <p>Aboriginal and Torres Strait Islander status for persons diagnosed in 2014 were impacted by a</p>	<p>1 – Aboriginal 2 – Torres Strait Islander 3 – Both Aboriginal and Torres Strait Islander 4 – Neither 8 – Declined 9 – Unknown</p>

Item No.	Element	Description/notes	Data values/format
		<p>system issue, producing an increased proportion of unknown Aboriginal and Torres Strait Islander status. The system issue is planned to be rectified for the next cancer incidence and mortality dataset release.</p> <p><u>Periods of use:</u> 1999 to latest available (for incidence) 1994 to latest available (for mortality)</p>	
4.	Year of birth	The year of birth of the person.	<p><i>Unknown</i> date of birth = 17 Nov 1858</p> <p>Format: YYYY</p>
5.	Month of birth	The month of birth of the person.	<p><i>Unknown</i> date of birth = 17 Nov 1858</p> <p>Format: MM</p>
6.	Day of birth	The day of birth of the person.	<p><i>Unknown</i> date of birth = 17 Nov 1858</p> <p>Format: DD</p>
7.	Date of birth validity code	<p>A code reflecting the accuracy of a person's date of birth.</p> <p>Date of birth was not a mandatory item prior to 1994</p>	<p>1 – Actual date (day, month, year known) 2 – Date unknown 4 – Day and month estimated (year known)</p>
Cancer diagnosis data elements			
8.	Year of diagnosis	<p>The year on which the person was first diagnosed with cancer.</p> <p>Date of diagnosis reflects the earliest definitive diagnostic or treatment episode received by the NSW Cancer Registry and may not be the exact date of diagnosis.</p> <p><u>Caveat</u> Please refer to caveat for further details.</p>	<p><i>Unknown</i> year of diagnosis = 1888</p> <p>Format: YYYY</p>
9.	Month of diagnosis	<p>The month on which the person was first diagnosed with cancer.</p> <p>Date of diagnosis reflects the earliest definitive diagnostic or treatment episode received by the NSW Cancer Registry and may not be the exact date of diagnosis.</p> <p><u>Caveat</u> Please refer to caveat for further details.</p>	<p><i>Unknown</i> month of diagnosis = Nov</p> <p>Format: MM</p>
10.	Day of diagnosis	<p>The day on which the person was first diagnosed with cancer.</p> <p>Date of diagnosis reflects the earliest definitive diagnostic or treatment episode received by the</p>	<p><i>Unknown</i> day of diagnosis = 11th</p> <p>Format: DD</p>

Item No.	Element	Description/notes	Data values/format
		NSWCR and may not be the exact date of diagnosis. <u>Caveat</u> Please refer to caveat for further details.	
11.	Date of diagnosis validity code	A code reflecting the accuracy of a date of diagnosis.	1 – Actual date (day, month, year known) 2 – Date unknown 3 – Day estimated (month, year known) 4 – Day and month estimated (year known)
12.	Age at diagnosis	The age at which the person was first diagnosed with cancer. Calculated from date of birth and date of diagnosis.	<i>Unknown age at diagnosis = Null</i> Format: NNN
13.	Cancer type	The type of cancer according to the reporting categories for cancer types. The tabulation category is based on ICD-9 and is derived from ICD-O-3 topography and morphology.	<i>See Attachment 1 – Cancer type reporting categories</i> Format: CNNN
14.	Clinical cancer group	The clinical cancer group of the cancer at diagnosis. Clinical cancer groups are a grouping of cancers based on treatment categories developed with advice from Professor Jim Bishop and endorsed by the Australasian Association of Cancer Registries (AACR) in December 2004. Clinical group is derived from the grouping of cancer types defined in element 13.	<i>See Attachment 2 – Clinical cancer group</i> Format: CNN
15.	Topography code (ICD-O-3)	The anatomical site of origin (or organ system) of the cancer at diagnosis. Coded according to the WHO International Classification of Diseases for Oncology (third edition).	http://www.who.int/classifications/icd/adaptations/oncology/en/ Format: CNNN
16.	Topography code (ICD-10-AM)	The anatomical site of origin (or organ system) of the cancer at diagnosis. Coded according to the 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.	http://sydney.edu.au/health-sciences/ncch/icd-10-am.shtml Format: CNNN
17.	Morphology code (ICD-O-3)	The morphology code of the cancer at diagnosis. Morphology codes describe the cell type of the cancer according to the WHO International Classification of Diseases for Oncology (third edition). A three digit morphology code refers to the main grouping by cell type e.g. Lymphoid leukaemia.	http://www.who.int/classifications/icd/adaptations/oncology/en/ Format: NNN

Item No.	Element	Description/notes	Data values/format
		<p><u>Caveat</u> Morphology codes may change over time as classifications change. E.g. Myelodysplasia was previously considered to have uncertain tumour behaviour, but since 2001 this element has been considered an invasive cancer. Data element is available from 2003 year of diagnosis.</p>	
18.	Morphology code 4 digit (ICD-O-3)	<p>The morphology code of the cancer at diagnosis. Morphology codes describe the cell type of the cancer according to the WHO International Classification of Diseases for Oncology (third edition). A four digit morphology code contains a more specific morphology e.g. Acute lymphoblastic leukaemia, not otherwise specified.</p> <p><u>Caveat</u> Morphology codes may change over time as classifications change. E.g. Myelodysplasia was previously considered to have uncertain tumour behaviour, but since 2001 this element has been considered an invasive cancer. Data element is available from 2003 year of diagnosis.</p>	<p>http://www.who.int/classifications/icd/adaptations/oncology/en/</p> <p>Format: NNNN</p>
19.	Behaviour code	<p>The behaviour of the cancer at diagnosis.</p> <p>It is equivalent to the 5th digit of the ICD-O-3 morphology code.</p>	3 – Invasive
20.	Best basis of diagnosis	<p>Best basis of diagnosis or 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 NSW Cancer Registry (NSWCR) coder, then the method at diagnosis is 'Histopathology sighted at NSWCR'.</p>	<p>0 – Cytology including FNA, smears, washing, sputum 1 – Clinical/imaging/ biochemical 2 – Histopathology performed 5 – Death certificate only 6 – Histopathology sighted at NSWCR</p>
21.	Degree of spread at diagnosis	<p>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.</p> <p><u>Caveat</u> <i>Please refer to caveat document for further details.</i></p>	<p>1 – Localised to tissue of origin 2 – Regional spread, adjacent organs and/or regional lymph nodes 3 – Distant metastases 9 – Unknown</p>
22.	Laterality	<p>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.</p> <p>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).</p>	<p>L – left R – right</p>
23.	Breslow thickness of melanoma /	<p>Cutaneous melanoma thickness is recorded completely for all cases diagnosed in 1988 onwards.</p>	<p>Valid values are 0.1-994 (mm). Other values are: 995 = micro-invasion</p>

Item No.	Element	Description/notes	Data values/format
	Size of breast cancer	Thickness (mm) is the thickness of the total tissue extracted. For breast cancer , tumour 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.	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 primary cancer cases reported for statistical purposes in the NSW reporting data set. Cancer cases are subjected to the international rules for determining multiple primary cancers developed by the International Association of Cancer Registries.	Format: NN
25.	Registry derived-stage (STaR)	Simplified TNM stage at diagnosis derived from data sources available to the NSW Cancer Registry and collected as part of Cancer Australia's Stage, Treatment and Recurrence (STaR) project. Cancer Australia has supported the Cancer Council Victoria to develop business rules to derive TNM stage at diagnosis based on data sources available to all population-based cancer registries. The rules were endorsed by the Australasian Association of Cancer Registries (AACR). Registry derived stage is available for 2011 year of diagnosis only, for the five highest incidence cancers (prostate, breast, lung, bowel and melanoma),	1 – Stage I 2 – Stage II 3 – Stage III 4 – Stage IV 88 – Stage not applicable 9 – Stage missing/unavailable (includes morphologies not eligible for TNM staging and tumours with no histological confirmation (as per AJCC rules)) <u>Cancer Australia:</u> https://canceraustralia.gov.au/research-data/cancer-data/improving-cancer-data
26.	Registry derived staging basis (STaR)	Accompanies registry derived-stage to clarify the source of the information used in the derivation of stage for a particular tumour.	P – pathology only C – clinical only I – integrated (combined clinical and pathological stage) U – unknown
Mortality data elements			
27.	Year of death	The year of death of the person	Format: YYYY
28.	Month of death	The month of death of the person	Format: MM
29.	Day of death	The day of death of the person	Format: DD
30.	Age at death	The age of the person at death	Format: NNN
31.	Cause of death cancer type	The cancer type that has led to the death of the person, according to the reporting categories for cancer types. The tabulation category is based on ICD-9 and is derived from the ICD-O-3 topography and morphology that has caused the death.	See <i>Attachment 1</i> – Cancer type reporting categories Format: CNNN
32.	Cause of death clinical cancer group	The clinical cancer group that has led to the death of the person. Clinical cancer groups are a grouping of cancers based on treatment categories developed with advice from Professor Jim Bishop	See <i>Attachment 2</i> – Clinical cancer group Format: CNN

Item No.	Element	Description/notes	Data values/format
		and endorsed by the Australasian Association of Cancer Registries (AACR) in December 2004. Cause of death clinical cancer group is derived from the grouping of cause of death cancer types in Item 31.	
33.	Cause of death topography code (ICD-O-3)	The anatomical site of origin (or organ system) of the cancer that has caused the death of the person. Coded according to the WHO International Classification of Diseases for Oncology (third edition).	http://www.who.int/classifications/icd/adaptations/oncology/en/ Format: CNNN
34.	Cause of death topography code (ICD-10-AM)	The anatomical site of origin (or organ system) of the cancer that has caused the death of the person. Coded according to the 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.	http://sydney.edu.au/health-sciences/ncch/icd-10-am.shtml Format: CNNN
35.	Place of death group	Persons that have been diagnosed with cancer in NSW and have subsequently died of cancer in NSW have been grouped into specific categories of place of death as approved by the NSW Registry of Births Deaths and Marriages. <u>Caveat</u> Please refer to caveat document for further details.	1 – Died at Home 2 – Died in a Hospice 3 – Died in a Public Hospital 4 – Died in a Private Hospital 5 – Died in a Nursing Home 9 – Unknown Place of Death NA – Not applicable
Geographical data elements (based on residence at diagnosis)			
36.	Postcode	Postcode of residence at diagnosis	Format: NNNN
37.	LGA 2006 (ASGC)	Local government area (LGA) of residence at diagnosis The LGA is a spatial unit which represents the whole geographical area of responsibility of an incorporated, officially gazetted local government area/council, as defined by each state and territory. It is based on the Australian Standard Geographical Classification (ASGC 2006), issued by the Australian Bureau of Statistics (ABS), effective at 1 st July 2006.	http://www.abs.gov.au/AUSSTATS/abs@.nsf/Previousproducts/943E92C473A3C463CA2571A900177C33?opendocument
38.	SLA 2006 (ASGC)	Statistical local area (SLA) of residence at diagnosis The SLA is a general purpose spatial unit that covers the whole of Australia without gaps or overlaps. It is based on the Australian Standard Geographical Classification (ASGC 2006), issued by	http://www.abs.gov.au/AUSSTATS/abs@.nsf/Previousproducts/907A694590693B32CA2571A9001784A4?opendocument

Item No.	Element	Description/notes	Data values/format												
		the Australian Bureau of Statistics (ABS), effective at 1 st July 2006.													
39.	Remoteness 2006 (ASGC)	<p>Remoteness of residence at diagnosis</p> <p>Remoteness areas divide Australia into 5 classes of remoteness on the basis of a measure of relative access to services. Access to services are measures using the Accessibility/Remoteness Index of Australia (ARIA+) methodology developed by the National Key Centre for Social Applications of GIS (GISCA) and utilised by the Australian Bureau of Statistics (ABS). GISCA is now incorporated by the University of Adelaide's Hugo Centre for Migration and Population Research. It is based on the Australian Standard Geographical Classification (ASGC) 2006 edition.</p> <p>Remoteness is available from year of diagnosis 2000 and onwards.</p>	<p>0 – Major cities 1 – Inner regional 2 – Outer regional 3 – Remote 4 – Very remote 9 – Unknown</p> <p><u>Remoteness 2006</u> http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1259.0.30.004Main+Features12006</p>												
40.	Socioeconomic position – IRSD quintiles (ASGC)	<p>Socioeconomic position of residence at diagnosis.</p> <p>Socioeconomic position is a general socio-economic index that summarises a range of information about the economic and social conditions of people and households within a geographic area. The index of relative socioeconomic disadvantage (IRSD) is developed by the Australian Bureau of Statistics (ABS). It is based on the Australian Standard Geographical Classification edition (ASGC) and the edition applied depends on the year of diagnosis:</p> <table border="0"> <tr> <td><u>Year of diagnosis</u></td> <td><u>Methodology used</u></td> </tr> <tr> <td>1980 to 1988</td> <td>CD (ASGC 1986)</td> </tr> <tr> <td>1989 to 1993</td> <td>CD (ASGC 1991)</td> </tr> <tr> <td>1994 to 1998</td> <td>CD (ASGC 1996)</td> </tr> <tr> <td>1999 to 2003</td> <td>CD (ASGC 2001)</td> </tr> <tr> <td>2004 onwards</td> <td>CD (ASGC 2006)</td> </tr> </table> <p>Socioeconomic position for IRSD is available from year of diagnosis 1980 and onwards, using the relevant census year index.</p> <p><u>Caveat</u> The ABS indicates that the census year indexes are not designed for analysis over time and should not be attempted lightly. For more information please refer to the ABS website: http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifahelpansuis?opendocument&navpos=260</p>	<u>Year of diagnosis</u>	<u>Methodology used</u>	1980 to 1988	CD (ASGC 1986)	1989 to 1993	CD (ASGC 1991)	1994 to 1998	CD (ASGC 1996)	1999 to 2003	CD (ASGC 2001)	2004 onwards	CD (ASGC 2006)	<p>1 – Quintile 1 (most disadvantaged) 2 – Quintile 2 3 – Quintile 3 4 – Quintile 4 5 – Quintile 5 (least disadvantaged)</p> <p><u>Socioeconomic position (2006)</u> http://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0</p>
<u>Year of diagnosis</u>	<u>Methodology used</u>														
1980 to 1988	CD (ASGC 1986)														
1989 to 1993	CD (ASGC 1991)														
1994 to 1998	CD (ASGC 1996)														
1999 to 2003	CD (ASGC 2001)														
2004 onwards	CD (ASGC 2006)														
41.	LGA 2016 (ASGS)	<p>Local government area (LGA) of residence at diagnosis</p> <p>The LGA is a spatial unit which represents the whole geographical area of responsibility of an incorporated, officially gazetted local government</p>	<p>http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.003~July%202016~Main%20Features~Local%20Government%20Areas%20(LGA)~7</p>												

Item No.	Element	Description/notes	Data values/format
		area or council, as defined by each state and territory. It is based on the Australian Statistical Geography Standard (ASGS 2016), issued by the Australian Bureau of Statistics (ABS), effective at 1 st July 2016.	
42.	SA2 2016 (ASGS)	<p>Statistical area level 2 (SA2) of residence at diagnosis</p> <p>The SA2 is a general-purpose medium-sized spatial unit built from whole SA1s. Their aim is to represent a community that interacts together socially and economically. It is based on the Australian Statistical Geography Standard (ASGS 2016), issued by the Australian Bureau of Statistics (ABS), effective at 1st July 2016.</p>	http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Related%20material%20and%20release%20timetable~10001
43.	SA3 2016 (ASGS)	<p>Statistical area level 3 (SA3) of residence at diagnosis</p> <p>The SA3 is a regional spatial unit built from whole SA2s. Their aim is to create a standard framework for the analysis of Australian Bureau of Statistics (ABS) data at the regional level through grouping SA2s with similar regional characteristics. It is based on the Australian Statistical Geography Standard (ASGS 2016), issued by the ABS, effective at 1st July 2016.</p>	http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Related%20material%20and%20release%20timetable~10001
44.	SA4 2016 (ASGS)	<p>Statistical area level 4 (SA4) of residence at diagnosis</p> <p>The SA4 is a standardised spatial unit built from whole SA3s. They are designed for the output of labour force data and reflect labour markets within each state and territory. It is based on the Australian Statistical Geography Standard (ASGS 2016), issued by the Australian Bureau of Statistics (ABS), effective at 1st July 2016.</p>	http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Related%20material%20and%20release%20timetable~10001
45.	GCCSA 2016 (ASGS)	<p>Greater capital city statistical area (GCCSA) of residence at diagnosis</p> <p>The GCCSAs represent the socio-economic extent of state and territory capital cities and are built up from whole SA4s. It is based on the Australian Statistical Geography Standard (ASGS 2016), issued by the Australian Bureau of Statistics (ABS), effective at 1st July 2016.</p>	http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~Main%20Features~Related%20material%20and%20release%20timetable~10001
46.	Remoteness (ASGC, ASGS)	<p>Remoteness of residence at diagnosis</p> <p>Remoteness is based on the Accessibility/Remoteness Index of Australia (ARIA+) methodology developed by the National Key Centre for Social Applications of GIS (GISCA) and utilised by the Australian Bureau of Statistics (ABS). GISCA is now incorporated by the University of Adelaide's Hugo Centre for Migration and Population Research. It is based on a combination</p>	<p>0 – Major cities 1 – Inner regional 2 – Outer regional 3 – Remote 4 – Very remote -1 – Not available (prior to year of diagnosis 2000) -2 – No score available (valid boundary)</p>

Item No.	Element	Description/notes	Data values/format										
		<p>of the Australian Standard Geographical Classification (ASGC) and the Australian Statistical Geography Standard (ASGC), depending on the years of diagnosis:</p> <table border="0"> <tr> <td><u>Year of diagnosis</u></td> <td><u>Methodology used</u></td> </tr> <tr> <td>2000 to 2010</td> <td>CD (ASGC 2006)</td> </tr> <tr> <td>2011 to 2013</td> <td>SA1 (ASGS 2011)</td> </tr> <tr> <td>2014 to 2018</td> <td>SA1 (ASGS 2016)</td> </tr> </table> <p>Remoteness is assigned to records pertaining to year of diagnosis 2000 and onwards.</p>	<u>Year of diagnosis</u>	<u>Methodology used</u>	2000 to 2010	CD (ASGC 2006)	2011 to 2013	SA1 (ASGS 2011)	2014 to 2018	SA1 (ASGS 2016)	<p>-3 – Unknown boundary (unknown address)</p> <p><u>Remoteness 2016 (ASGS)</u> http://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005</p>		
<u>Year of diagnosis</u>	<u>Methodology used</u>												
2000 to 2010	CD (ASGC 2006)												
2011 to 2013	SA1 (ASGS 2011)												
2014 to 2018	SA1 (ASGS 2016)												
47.	Remoteness calculation method	Describes which ABS classification system and which geographical area has been used to calculate remoteness	<p>-1 – Not available (prior to year of diagnosis 2000)</p> <p>Format: YYYY-AAN</p>										
48.	Socioeconomic position - IRSAD deciles (ASGC, ASGS)	<p>Socioeconomic position of residence at diagnosis.</p> <p>Socioeconomic position is a general socio-economic index that summarises a range of information about the economic and social conditions of people and households within a geographic area. The index of relative socioeconomic advantage and disadvantage (IRSAD) is developed by the Australian Bureau of Statistics (ABS). It is based on a combination of the Australian Standard Geographical Classification (ASGC) and the Australian Statistical Geography Standard (ASGC), depending on the years of diagnosis:</p> <table border="0"> <tr> <td><u>Year of diagnosis</u></td> <td><u>Methodology used</u></td> </tr> <tr> <td>1999 to 2003</td> <td>CD (ASGC 2001)</td> </tr> <tr> <td>2004 to 2010</td> <td>CD (ASGC 2006)</td> </tr> <tr> <td>2011 to 2013</td> <td>SA1 (ASGS 2011)</td> </tr> <tr> <td>2014 to 2018</td> <td>SA1 (ASGS 2016)</td> </tr> </table> <p>Socioeconomic position for IRSAD is assigned to records for year of diagnosis 1999 and onwards, using the relevant census year index.</p> <p><u>Caveat</u> The ABS indicates that the census year indexes are not designed for analysis over time and should not be attempted lightly. For more information please refer to the ABS website: http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifahelpansuis?opendocument&navpos=260</p>	<u>Year of diagnosis</u>	<u>Methodology used</u>	1999 to 2003	CD (ASGC 2001)	2004 to 2010	CD (ASGC 2006)	2011 to 2013	SA1 (ASGS 2011)	2014 to 2018	SA1 (ASGS 2016)	<p>1 – Decile 1 (most disadvantaged) 2 – Decile 2 3 – Decile 3 4 – Decile 4 5 – Decile 5 6 – Decile 6 7 – Decile 7 8 – Decile 8 9 – Decile 9 10 – Decile 10 (least disadvantaged) -1 – Not available (prior to year of diagnosis 1980) - 2 – No score available (valid boundary but no score for year of diagnosis 1980 to 1998) -3 – Unknown boundary (unknown address)</p> <p><u>Socioeconomic position (2016)</u> http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2016~Main%20Features~SEIFA%20Basics~5</p>
<u>Year of diagnosis</u>	<u>Methodology used</u>												
1999 to 2003	CD (ASGC 2001)												
2004 to 2010	CD (ASGC 2006)												
2011 to 2013	SA1 (ASGS 2011)												
2014 to 2018	SA1 (ASGS 2016)												
49.	Socioeconomic position - IRSAD quintiles (ASGC, ASGS)	<p>Socioeconomic position of residence at diagnosis.</p> <p>Socioeconomic position is a general socio-economic index that summarises a range of information about the economic and social conditions of people and households within a geographic area. The index of relative socioeconomic advantage and disadvantage (IRSAD), developed by the Australian Bureau of</p>	<p>1 – Quintile 1 (most disadvantaged) 2 – Quintile 2 3 – Quintile 3 4 – Quintile 4 5 – Quintile 5 (least disadvantaged) -1 – Not available (prior to year of diagnosis 1980)</p>										

Item No.	Element	Description/notes	Data values/format																
		<p>Statistics (ABS). It is based on a combination of the Australian Standard Geographical Classification (ASGC) and the Australian Statistical Geography Standard (ASGC), depending on the years of diagnosis:</p> <table border="0"> <tr> <td><u>Year of diagnosis</u></td> <td><u>Methodology used</u></td> </tr> <tr> <td>1999 to 2003</td> <td>CD (ASGC 2001)</td> </tr> <tr> <td>2004 to 2010</td> <td>CD (ASGC 2006)</td> </tr> <tr> <td>2011 to 2013</td> <td>SA1 (ASGS 2011)</td> </tr> <tr> <td>2014 to 2018</td> <td>SA1 (ASGS 2016)</td> </tr> </table> <p>Socioeconomic position for IRSAD is assigned to records for year of diagnosis 1999 and onwards, using the relevant census year index.</p> <p><u>Caveat</u> The ABS indicates that the census year indexes are not designed for analysis over time and should not be attempted lightly. For more information please refer to the ABS website: http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifahelpansuis?opendocument&navpos=260</p>	<u>Year of diagnosis</u>	<u>Methodology used</u>	1999 to 2003	CD (ASGC 2001)	2004 to 2010	CD (ASGC 2006)	2011 to 2013	SA1 (ASGS 2011)	2014 to 2018	SA1 (ASGS 2016)	<p>-2 – (valid boundary but no score for year of diagnosis 1980 to 1998) -3 – Unknown boundary (unknown address)</p> <p><u>Socioeconomic position (2016)</u> http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2016~Main%20Features~SEIFA%20Basics~5</p>						
<u>Year of diagnosis</u>	<u>Methodology used</u>																		
1999 to 2003	CD (ASGC 2001)																		
2004 to 2010	CD (ASGC 2006)																		
2011 to 2013	SA1 (ASGS 2011)																		
2014 to 2018	SA1 (ASGS 2016)																		
50.	Socioeconomic position – IRSD deciles (ASGC, ASGS)	<p>Socioeconomic position of residence at diagnosis.</p> <p>Socioeconomic position is a general socio-economic index that summarises a range of information about the economic and social conditions of people and households within a geographic area. The index of relative socioeconomic disadvantage (IRSD) is developed by the Australian Bureau of Statistics (ABS). It is based on a combination of the Australian Standard Geographical Classification (ASGC) and the Australian Statistical Geography Standard (ASGC), depending on the years of diagnosis:</p> <table border="0"> <tr> <td><u>Year of diagnosis</u></td> <td><u>Methodology used</u></td> </tr> <tr> <td>1980 to 1988</td> <td>CD (ASGC 1986)</td> </tr> <tr> <td>1989 to 1993</td> <td>CD (ASGC 1991)</td> </tr> <tr> <td>1994 to 1998</td> <td>CD (ASGC 1996)</td> </tr> <tr> <td>1999 to 2003</td> <td>CD (ASGC 2001)</td> </tr> <tr> <td>2004 to 2010</td> <td>CD (ASGC 2006)</td> </tr> <tr> <td>2011 to 2013</td> <td>SA1 (ASGS 2011)</td> </tr> <tr> <td>2014 to 2018</td> <td>SA1 (ASGS 2016)</td> </tr> </table> <p>Socioeconomic position for IRSD is assigned to records for year of diagnosis 1980 and onwards, using the relevant census year index.</p> <p><u>Caveat</u> The ABS indicates that the census year indexes are not designed for analysis over time and should not be attempted lightly. For more information please refer to the ABS website:</p>	<u>Year of diagnosis</u>	<u>Methodology used</u>	1980 to 1988	CD (ASGC 1986)	1989 to 1993	CD (ASGC 1991)	1994 to 1998	CD (ASGC 1996)	1999 to 2003	CD (ASGC 2001)	2004 to 2010	CD (ASGC 2006)	2011 to 2013	SA1 (ASGS 2011)	2014 to 2018	SA1 (ASGS 2016)	<p>1 – Decile 1 (most disadvantaged) 2 – Decile 2 3 – Decile 3 4 – Decile 4 5 – Decile 5 6 – Decile 6 7 – Decile 7 8 – Decile 8 9 – Decile 9 10 – Decile 10 (least disadvantaged) -1 – Not available (prior to year of diagnosis 1980) -2 – No score available (valid boundary) -3 – Unknown boundary (unknown address)</p> <p><u>Socioeconomic position (2016)</u> http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2016~Main%20Features~SEIFA%20Basics~5</p>
<u>Year of diagnosis</u>	<u>Methodology used</u>																		
1980 to 1988	CD (ASGC 1986)																		
1989 to 1993	CD (ASGC 1991)																		
1994 to 1998	CD (ASGC 1996)																		
1999 to 2003	CD (ASGC 2001)																		
2004 to 2010	CD (ASGC 2006)																		
2011 to 2013	SA1 (ASGS 2011)																		
2014 to 2018	SA1 (ASGS 2016)																		

Item No.	Element	Description/notes	Data values/format																
		http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifahelpansuis?opendocument&navpos=260																	
51.	Socioeconomic position – IRSD quintiles (ASGC, ASGS)	<p>Socioeconomic position of residence at diagnosis.</p> <p>Socioeconomic position is a general socio-economic index that summarises a range of information about the economic and social conditions of people and households within a geographic area. The index of relative socioeconomic disadvantage (IRSD) is developed by the Australian Bureau of Statistics (ABS). It is based on a combination of the Australian Standard Geographical Classification (ASGC) and the Australian Statistical Geography Standard (ASGS), depending on the years of diagnosis:</p> <table border="0"> <tr> <td><u>Year of diagnosis</u></td> <td><u>Methodology used</u></td> </tr> <tr> <td>1980 to 1988</td> <td>CD (ASGC 1986)</td> </tr> <tr> <td>1989 to 1993</td> <td>CD (ASGC 1991)</td> </tr> <tr> <td>1994 to 1998</td> <td>CD (ASGC 1996)</td> </tr> <tr> <td>1999 to 2003</td> <td>CD (ASGC 2001)</td> </tr> <tr> <td>2004 to 2010</td> <td>CD (ASGC 2006)</td> </tr> <tr> <td>2011 to 2013</td> <td>SA1 (ASGS 2011)</td> </tr> <tr> <td>2014 to 2018</td> <td>SA1 (ASGS 2016)</td> </tr> </table> <p>Socioeconomic position for IRSD is assigned to records for year of diagnosis 1980 and onwards, using the relevant census year index.</p> <p><u>Caveat</u> The ABS indicates that the census year indexes are not designed for analysis over time and should not be attempted lightly. For more information please refer to the ABS website: http://www.abs.gov.au/websitedbs/censushome.nsf/home/seifahelpansuis?opendocument&navpos=260</p>	<u>Year of diagnosis</u>	<u>Methodology used</u>	1980 to 1988	CD (ASGC 1986)	1989 to 1993	CD (ASGC 1991)	1994 to 1998	CD (ASGC 1996)	1999 to 2003	CD (ASGC 2001)	2004 to 2010	CD (ASGC 2006)	2011 to 2013	SA1 (ASGS 2011)	2014 to 2018	SA1 (ASGS 2016)	<p>1 – Quintile 1 (most disadvantaged) 2 – Quintile 2 3 – Quintile 3 4 – Quintile 4 5 – Quintile 5 (least disadvantaged) -1 – Not available (prior to year of diagnosis 1980) -2 – No score available -3 – Unknown boundary</p> <p><u>Socioeconomic position (2016)</u> http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2016~Main%20Features~SEIFA%20Basics~5</p>
<u>Year of diagnosis</u>	<u>Methodology used</u>																		
1980 to 1988	CD (ASGC 1986)																		
1989 to 1993	CD (ASGC 1991)																		
1994 to 1998	CD (ASGC 1996)																		
1999 to 2003	CD (ASGC 2001)																		
2004 to 2010	CD (ASGC 2006)																		
2011 to 2013	SA1 (ASGS 2011)																		
2014 to 2018	SA1 (ASGS 2016)																		
52.	Socioeconomic position calculation method	Describes which ABS classification system and which geographical area has been used to calculate socioeconomic position	<p>-1 – Not available (prior to year of diagnosis 1980)</p> <p>Format: YYYY-AAAN</p>																
53.	Local health district	<p>The local health district is a 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.</p> <p>Local health district (LHD) is determined from residence at diagnosis.</p>	http://www.health.nsw.gov.au/lhd/pages/default.aspx																
54.	Primary health network	The primary health network as defined by NSW Health, in July 2015. Ten primary health networks cover the Sydney metropolitan region, rural and regional NSW.	http://www.health.gov.au/internet/main/publishing.nsf/Content/PHN-Home																

Item No.	Element	Description/notes	Data values/format
		Primary health network (PHN) is determined from residence at diagnosis.	
Episode of care data elements - these data elements are not population-based and have undergone minimal data validation and quality checks. Data is currently available for episodes of care that began in 2013.			
55.	Data source type	The source of the information for the episode of care.	I – Inpatient C – Medical Oncology O – Outpatient P – Pathology R – Radiotherapy
56.	Episode modality	The modality of the episode of care.	1 – Surgery 2 – Radiotherapy 3 – Medical oncology 4 – Admitted/other 5 – Diagnostic 6 – Terminal palliation
57.	Episode start date	The date the episode of care started.	Format: DDMMYYYY
58.	Episode end date	The date the episode of care ended. End dates for medical oncology treatment may not be available for oral and hormonal therapies that may be taken over a period of years.	Format: DDMMYYYY <i>Unknown year = 9999</i>
59.	Facility name	The facility where the episode of care occurred.	Episodes of care occurring in facilities not able to be identified are grouped together as the following codes: Interstate facility Private facility Unknown facility
60.	Local health district of Facility	Local health district of facility	X630 Sydney Children’s Hospitals Network X690 St Vincent’s Health Network X700 Sydney LHD X710 South Western Sydney LHD X720 South Eastern Sydney LHD X730 Illawarra Shoalhaven LHD X740 Western Sydney LHD X750 Nepean Blue Mountains LHD X760 North Sydney LHD X770 Central Coast LHD X800 Hunter New England LHD X810 Northern NSW LHD X820 Mid North Coast LHD X830 Southern NSW LHD X840 Murrumbidgee LHD X850 Western NSW LHD X860 Far West LHD UNKN LHD unknown Episodes of care occurring in facilities not able to be identified are grouped together with the following codes: INTR Interstate facility

Item No.	Element	Description/notes	Data values/format
			PRIV Private facility UNKN Unknown facility Format: ANNN
61.	Degree of spread at episode	The extent of spread of cancer from its point of origin at each episode of care.	1 – Localised to tissue of origin 2 – Regional spread, adjacent organs and/or regional lymph nodes 3 – Distant metastases 9 – Unknown
62.	TNM staging group	<p>The overall TNM stage value describes the extent of an individual's cancer spread based on the primary tumour and the extent of spread in the body, according to the clinically determined values of T, N and M.</p> <p>T stage represents the size and extent of the primary tumour; N stage represents the extent of regional lymph node metastasis; and M stage identifies the presence or absence of distant metastasis.</p> <p>The TNM staging system is an internationally agreed staging classification system developed by the American Joint Committee on Cancer (AJCC) and adopted by the Union for International Cancer Control (UICC).</p> <p>TNM staging applies to most solid tumours, but is not applicable to all tumours.</p>	4 – I 6 – IA 7 – IA1 8 – IA2 9 – IB 10 – IB1 11 – IB2 12 – IC 13 – II 14 – IIA 15 – IIA1 16 – IIA2 17 – IIB 18 – IIC 19 – III 20 – IIIA 21 – IIIB 22 – IIIC 30 – IIIC1 31 – IIIC2 23 – IV 24 – IVA 25 – IVA1 26 – IVA2 27 – IVB 28 – IVC 29 – Occult Carcinoma 97 – Unstaged 98 – Not Applicable 99 – Unknown
63.	TNM edition	TNM edition number	1 – TNM edition 1 2 – TNM edition 2 3 – TNM edition 3 4 – TNM edition 4 5 – TNM edition 5 6 – TNM edition 6 7 – TNM edition 7 9 – Unknown
64.	TNM staging basis	<p>Describes the evidence for the T, N and M stage values and the overall TNM stage value.</p> <p>Clinical stage is based on evidence obtained prior to treatment from physical examination, imaging, endoscopy, biopsy, surgical exploration or other relevant examinations.</p> <p>Pathological stage is based on histological evidence acquired before treatment, and is supplemented or</p>	C – Clinical P – Pathological

Item No.	Element	Description/notes	Data values/format
		<p>modified by additional evidence acquired from surgery and from pathological examination.</p> <p>When more than one basis for staging is available, the highest stage in any one of the stages is used.</p>	
65.	TNM staging timing	<p>An indicator of whether the TNM staging occurred at diagnosis or not at diagnosis.</p> <p>Staging at diagnosis is defined within four months of the date of diagnosis.</p>	<p>1 – Staging at diagnosis 2 – Staging not at diagnosis</p>
66.	TNM staging date	Date on which TNM staging was determined.	Format: DDMMYYYY
67.	MDT date	<p>Date of consultation with a multidisciplinary team of health professionals regarding the diagnosis, treatment and care of the person with cancer.</p> <p>An MDT meeting in this context is a face-to-face meeting (or via video/teleconference) held at a defined time and place for the express purpose of discussing cases and deciding treatment recommendations.</p>	Format: DDMMYYYY
68.	Performance status (ECOG)	<p>A score given that appraises an individual's ability to manage activities of daily living at the time of the episode of care.</p> <p>Code set for performance status scale by USA Eastern Cooperative Oncology Group (ECOG). (Isselbacher et al, 1994)</p>	<p>0 – Fully active, able to carry on all pre-disease performance without restriction 1 – Restricted in physically strenuous activity but ambulatory, can perform light / sedentary work 2 – Ambulatory and self-caring, cannot work. Up and about > 50% of waking hours 3 – Only limited self-care, confined to bed or chair > 50% of waking hours 4 – Completely disabled, cannot self-care. Totally confined to bed or chair 9 – Unknown</p>

Attachment 1 – Cancer type reporting categories

Code	Description
C00	Lip
C012	Tongue
C036	Mouth
C078	Salivary glands
C0910	Oropharynx
C11	Nasopharynx
C123	Hypopharynx
C14	Other oral cavity and 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

Code	Description
C545	Uterus, body and 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	Ill-defined and unspecified site
M95	Other lymphatic haematopoietic
0006	Non cancer death
0007	Unknown cause of death

Attachment 2 – Clinical cancer group

Clinical Cancer Group	Cancer Type (ICD-O-3 codes)
Bone and connective tissue	Bone (C40,C41) Connective tissue, peripheral nerves (C47, C49)
Bowel	Colon (C18) Rectal (C19-C21)
Breast	Breast (C50)
Cancer unknown primary	Other and Ill-defined digestive organs (C26) Other and Ill-defined sites in the respiratory system and intrathoracic organs (C39) Retroperitoneum and peritoneum (C48) Other and Ill-defined sites (C76) Unknown primary site (C80)
Eye	Eye (C69)
Gynaecological	Cervical (C53) Uterine (C54, C55) Ovarian (C56, C57.0-C57.7) Placental (C58) Other female genital (C51, C52, C57.8-C57.9)
Head and neck	Tongue (C01, C02) Mouth (C03-C06) Salivary gland (C07, C08) Oropharyngeal (C09, C10) Nasopharyngeal (C11) Hypopharyngeal (C12, C13) Other oral cavity and pharyngeal (C14) Nose, sinuses, etc. (C30, C31) Laryngeal (C32)
Lymphohaematopoietic	Hodgkin’s disease (M965-M966) Non-Hodgkin’s lymphoma (M959, M967-M972, M974) Multiple myeloma (M973, M976) Acute lymphoblastic leukaemia (M9821) Other lymphoid leukaemia (M9820, M9822-M9827, M994) Acute myeloid leukaemia (M9861) Other myeloid leukaemia (M9860, M9862-8, M987-M988, M9930, M9987) Other specified leukaemia (M984, M985, M989-M993) Unspecified leukaemia (M980) Myelodysplasia (M998) Other lymphoid haematopoietic (C42,C77 and M974, M975, M995-M997 [excl. M9963])
Neurological	Brain (C71) Central nervous system (C72)
Respiratory	Lung (C33, C34) Other thoracic (C37, C38) Mesothelioma (M905)

Clinical Cancer Group	Cancer Type (ICD-O-3 codes)
Skin	Lip (C00)
	Melanoma of the skin (C44 and M872-M879)
	Kaposi's sarcoma (M914)
Thyroid and other endocrine	Thyroid (C73)
	Other endocrine glands (C74, C75)
Upper gastrointestinal	Oesophageal (C15)
	Stomach (C16)
	Small intestinal (C17)
	Liver (C22)
	Gallbladder (C23, C24)
	Pancreatic (C25)
Urogenital	Prostate (C61)
	Testicular (C62)
	Other male genital (C60, C63)
	Kidney (C64-C66, C68)
	Bladder (C67)