

**AIHW EO2013-1-7:  
The Australian Longitudinal Study on  
Women's Health**

**Australian Cancer Database  
Metadata (25<sup>th</sup> July 2018)**

## Contents

PROJECT_PERSON_NUMBER .....	3
SEX .....	3
STATE.....	3
BIRTH_DATE.....	4
BIRTH_DATE_ACCURACY .....	4
DIAGNOSIS_DATE.....	5
DIAGNOSIS_YEAR_MONTH.....	5
DIAGNOSIS_DATE_ACCURACY.....	5
DIAGNOSIS_AGE .....	5
DIAGNOSIS_AGE_GROUP.....	6
PROJECT_TUMOUR_NUMBER .....	6
TOPOGRAPHY .....	6
MORPHOLOGY.....	6
ICD10.....	6
BEST_BASIS_OF_DIAGNOSIS .....	7
MELANOMA_THICKNESS.....	8
BREAST_TUMOUR_SIZE.....	8
DEATH_DATE.....	8
DEATH_YEAR_MONTH .....	9
DEATH_AGE .....	9
DEATH_AGE_GROUP .....	9
UNDERLYING_CAUSE_OF_DEATH (ICD10).....	9
COUNTRY_OF_BIRTH.....	11
SLA_2006_AT_DIAGNOSIS.....	11
POSTCODE_AT_DIAGNOSIS.....	11

## PROJECT\_PERSON\_NUMBER

Type: Character, length 24.

Meaning: An alphanumeric code derived by AIHW which uniquely identifies the person in the project.

Meaning of null value: Impossible.

## SEX

Type: Numeric, integer.

Meaning: Sex of the person as recorded on the Australian Cancer Database (ACD).

Meaning of null value: Impossible.

Note: Four links were accepted in which the sex information recorded on the ACD was male ('1'). The source of this data was checked by identifying the links accepted between the ALSWH cohort and the ACD. All four links look correct as there was exact agreement on name information, date of birth and only the variable sex did not match. It is likely that there was clerical error when recording the sex of these four records on the ACD.

Value	Meaning
1	Male
2	Female

## STATE

Type: Character, length 3.

Meaning: State or territory of the usual residence of the person at the time of their diagnosis.

Meaning of null value: Impossible.

Value	Meaning
ACT	Australian Capital Territory
NSW	New South Wales or Jervis Bay Territory
NT	Northern Territory
QLD	Queensland
SA	South Australia
TAS	Tasmania
VIC	Victoria
WA	Western Australia, Christmas Island or Cocos (Keeling) Islands

## BIRTH\_DATE

Type: Date.

Meaning: Date on which the person was born.

Meaning of null value: Queensland and Tasmania have not approved the release of birth date information and so this field will be null for those states.

## BIRTH\_DATE\_ACCURACY

Type: Character, length 3.

Meaning: An indicator of the accuracy of birth\_date.

Meaning of null value: Queensland and Tasmania have not approved the release of birth date information and so this field will be null for those states.

Each accuracy indicator consists of three letters. Although the first letter generally relates only to the accuracy of the day, the second to the month and the third to the year, there are some cases (FFF and HH\*) in which the indicator means something about more than just one component. The following table explains the possibilities.

Value	Meaning
A (can occur in any position)	The component has been specifically recorded by the registry as accurate.
B (can occur in any position)	The registry does not (or did not at that time) have an accuracy indicator. Nevertheless, this component is likely to be accurate.
E (can occur in any position)	The true value of the component is unknown and has been estimated by the registry. Note that if one or both of the other components also have the value E, these components may have been estimated jointly rather than independently.
FFF	One or more components were estimated by the registry but it is not recorded which one(s).
G** (G in 1st position, other positions can be A, B or E)	The true value of 'day' is unknown and has been estimated by AIHW. It is set to the middle day of the month unless, in the case of diagnosis (i.e. not birth), the birth and/or death occurred in the same month and year as diagnosis, in which case the value of day of diagnosis is set to the middle day of the range of possible days in which diagnosis could have occurred (i.e. you can't be diagnosed before birth or after death).
HH* (HH in 1st and 2nd positions, 3rd position can be A, B or E)	The true values of day and month are unknown and have been set by AIHW to the middle day of the year, i.e. 2nd July.

## DIAGNOSIS\_DATE

Type: Date.

Meaning: Date on which the tumour was diagnosed.

Meaning of null value: New South Wales, Western Australia and the Australian Capital Territory have not approved the release of full diagnosis date information and so this field will be null for those states. However, each of these states and territories allowed the release of diagnosis\_year\_month.

## DIAGNOSIS\_YEAR\_MONTH

Type: Character, length 6.

Meaning: The year and month on which the tumour was diagnosed.

Meaning of null value: Impossible.

## DIAGNOSIS\_DATE\_ACCURACY

Type: Character, length 3.

Meaning: An indicator of the accuracy of diagnosis\_date.

Meaning of null value: Impossible.

Values: Same as those used for the field birth\_date\_accuracy.

## DIAGNOSIS\_AGE

Type: Numeric, integer.

Meaning: The age, in completed whole years, of the person on the day of their diagnosis.

Meaning of null value: Impossible.

Special value	Meaning
999	Unknown age at diagnosis

## **DIAGNOSIS\_AGE\_GROUP**

Type: Numeric, integer.

Meaning: The age group of the person on the day of their diagnosis.

Meaning of null value: Impossible.

Note: The age groups, numbered 1 to 18, are the standard age groups 0-4, 5-9, 10-14, ..., 80-84, 85+.

## **PROJECT\_TUMOUR\_NUMBER**

Type: Character, length 24.

Meaning: An alphanumeric code derived by AIHW which uniquely identifies the tumour in the project.

Meaning of null value: Impossible.

## **TOPOGRAPHY**

Type: Character, length 4.

Meaning: ICD-O-3 topography code of the tumour.

Meaning of null value: Impossible.

## **MORPHOLOGY**

Type: Character, length 6.

Meaning: The tumour histology concatenated with the tumour behaviour.

Meaning of null value: Impossible.

Note: The tumour behaviour is always '3' in this dataset - i.e. malignant tumours (cancers).

## **ICD10**

Type: Character, length 4.

Meaning: The ICD-10 (2016 version) disease code of the tumour.

Meaning of null value: Impossible.

## BEST\_BASIS\_OF\_DIAGNOSIS

Type: Numeric, integer.

Meaning: The most valid basis of diagnosis of the tumour.

Meaning of null value: Impossible.

Note: A tumour can be diagnosed by a number of different methods, each of which has a degree of reliability associated with it. The value contained in this field represents the method of the most reliable method used, using the hierarchy 7 > 6 > ... > 1 > 0 (see table below).

Value	Meaning
0	Death certificate only (DCO): information provided is from a death certificate. No other information has been found regarding the neoplasm
1	Clinical only: diagnosis made before death, but without any of the following (codes 2-8)
2	Clinical investigation: any diagnostic technique, including x-ray, endoscopy, imaging, ultrasound, exploratory surgery (e.g. laparotomy) and autopsy, without a tissue diagnosis
4	Specific tumour markers: including biochemical and/or immunological markers that are specific for a tumour site
5	Cytology: examination of cells from a primary or secondary site, including fluids aspirated by endoscopy or needle; also includes the microscopic examination of peripheral blood and bone marrow aspirates
6	Histology of metastasis: histological examination of tissue from a metastasis, including autopsy specimens
7	Histology of a primary tumour: histological examination of tissue from primary tumour, however obtained, including all cutting techniques and bone marrow biopsies; also includes autopsy specimens of primary tumour
8	Histology: either unknown whether of primary or metastatic site, or not specified
9	Unknown

## MELANOMA\_THICKNESS

Type: Numeric, real.

Meaning: The Breslow thickness of a malignant melanoma of the skin, measured in millimetres.

Meaning of null value: The tumour is not a malignant melanoma of the skin.

Special value	Meaning
Null	Tumour is not a malignant melanoma of the skin
9.89 (SA records only)	$\geq 9.89$
9.99 (VIC and NT records only)	$\geq 9.99$
990	Microinvasive
999	Unknown

## BREAST\_TUMOUR\_SIZE

Type: Numeric, real.

Meaning: The size of an invasive breast cancer, in millimetres.

Meaning of null value: The tumour is not an invasive breast cancer.

Special value	Meaning
Null	Tumour is not an invasive breast cancer
3 (SA records only)	$\leq 3$
990	Stated to be microinvasive but no size given
995	Occult primary (cancer cells were not found in the breast tissue excised but were found in at least one lymph node excised)
999	Unknown or not applicable (e.g. phyllodes tumour)

## DEATH\_DATE

Type: Date.

Meaning: Date on which the person died.

Meaning of null value: Person is still alive, or not known to be dead. The exception to this is occurs for the Australian Capital Territory in which did not agree to release the death\_date for those deceased, however they did agree to release the death\_year\_month.

## DEATH\_YEAR\_MONTH

Type: Character, length 6.

Meaning: Year and month on which the person died.

Meaning of null value: Person is still alive, or not known to be dead.

## DEATH\_AGE

Type: Numeric, integer.

Meaning: The age, in completed whole years, of the person at the time of their death.

Meaning of null value: Person is still alive, or not known to be dead.

Special value	Meaning
Null	Person is still alive, or not known to be dead
999	Unknown age at death

## DEATH\_AGE\_GROUP

Type: Numeric, integer.

Meaning: The age group of the person on the day of their death.

Meaning of null value: Person is still alive, or not known to be dead.

Note: The age groups, numbered 1 to 18, are the standard age groups 0-4, 5-9, ..., 80-84, 85+.

Special Value	Meaning
99	Unknown age at death

## UNDERLYING\_CAUSE\_OF\_DEATH

Type: Character, length 4.

Meaning: The ICD-10 code for the underlying cause of death, as assigned by the cancer registry.

Meaning of null value: Person is (1) still alive or (2) not known to be dead.

Note 1: The only codes used are the neoplasm codes, i.e. C000 to D489, and the special values in the table below. However, most cancer registries make little or no use of the non-

cancer neoplasm codes (D codes except for D45, D46, D47.1, D47.3–D47.5) and code such deaths to non-cancer deaths (ZNCD).

<b>Special value</b>	<b>Meaning</b>
Null	Person is (1) still alive or (2) not known to be dead.
Begins with C44	Rare non-melanoma skin cancer. Common NMSC is coded to Z44. Contrast this with <code>cause_of_death_ndi</code>
D03x, D05x, D06x, D32x, D33x, D42x, D43x	These codes are equivalent to their 3-character counterparts. The final x is used to signify that the 4th character has not been assigned. For example, D03x simply means D03, melanoma in situ.
Begins with Z44	Common non-melanoma skin cancer, i.e. BCC or SCC. Rare NMSC is coded to C44. Contrast this with <code>cause_of_death_ndi</code> .  Note: Most registries make little or no use of this code. I think NSW and ACT code such deaths to ZNCD.
ZBCC	Basal cell carcinoma of the skin
ZBEN	Benign neoplasm
ZBOR	Neoplasm of borderline malignancy (uncertain behaviour)
ZINS	In situ neoplasm
ZNCD	Non-cancer death.  Note: For some registries, this includes deaths from benign, uncertain behaviour and in situ neoplasms. For NSW and ACT I think it also includes deaths from common NMSC.
ZNDI	Person has been linked to the NDI by AIHW but the link has not yet been confirmed by the cancer registry. The cause of death recorded on the NDI is stored in the field <code>cause_of_death_ndi</code>
ZSCC	Squamous cell carcinoma of the skin
ZUNK	Unknown cause of death, i.e. could be cancer or non-cancer

## **COUNTRY\_OF\_BIRTH**

Type: Numeric, integer.

Meaning: Country in which the person was born.

Meaning of null value: All states and territories did not approve the release of Country of birth information excluding New South Wales and the Northern Territory.

Values and meanings are as per the Standard Australian Classification of Countries 2011.

Note 1: This field contains supplementary codes, i.e. codes in the range 0900–0999.

Note 2: Values which were permissible in earlier versions of the SACC but are not in the 2011 edition are permissible in this field. These values are 0002 (not elsewhere classified), 0923 (United Kingdom, nfd), 2101 (Channel Islands), 3213 (Serbia and Montenegro), 4199 (North Africa, nec) and 8418 (Netherlands Antilles).

Note 3: Due to a system issue in NSW, people diagnosed with cancer for the first time in 2011 or 2012 in NSW or 2011 to 2013 in the ACT were not being assigned the correct country of birth. As this issue had not been resolved by the time of data submission, all such people have been assigned an unknown country of birth. This means that we cannot do national country of birth analyses for 2011 onwards.

## **SLA\_2006\_AT\_DIAGNOSIS**

Type: Numeric, integer.

Meaning: 5-digit Statistical Local Area (SLA) of the usual residence of the person at the time of their diagnosis, in the 2006 ASGC scheme.

Meaning of null value: All states and territories did not approve the release of SLA information excluding New South Wales, Northern Territory and the Australian Capital Territory.

Note: If the 2006 SLA is unknown or the person has no usual address, sla2006 is set to the ABS state code followed by four 9s.

## **POSTCODE\_AT\_DIAGNOSIS**

Type: Numeric, integer.

Meaning: Postcode of the usual residence of the person at the time of their diagnosis.

Meaning of null value: All states and territories did not approve the release of postcode at diagnosis information excluding Western Australia.

Note: If the postcode is unknown or the person has no usual address, the following special values are used: 0899, 2699, 2999, 3999, 4999, 5999, 6999 and 7999.