



Government of **Western Australia**  
Department of **Health**

# **Synthetic WA Stomach Cancer Registry**

## Data Dictionary

**Western Australia Health**

**Important Disclaimer:**

The State of WA, the WA Department of Health and their respective officers, employees and agents, do not accept legal liability or responsibility for the Material, or any consequences arising from its use.

<b>Owner:</b>	Department of Health, Western Australia
<b>Contact:</b>	Information and Performance Governance Unit
<b>Approved by:</b>	Hamza Cakan
<b>Links to:</b>	Information Management Policy Framework <a href="https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management">https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management</a>

# Contents

<b>Abbreviations</b> .....	<b>4</b>
<b>Purpose</b> .....	<b>5</b>
<b>Background</b> .....	<b>5</b>
<b>Generation of data</b> .....	<b>5</b>
<b>Data definitions</b> .....	<b>5</b>
Aboriginal Status .....	6
Age Range .....	7
Age Range at death .....	7
Cancer type .....	8
Cause of death - cancer related .....	8
Cause of death - due to tumour .....	9
Detection method .....	9
Diagnosis Postcode Region .....	10
Grade .....	11
Laterality .....	13
Person Identifier .....	13
Recurrence - date of onset .....	14
Region of Birth .....	14
Sex .....	15
Tumour Site - Code .....	15
Year of Diagnosis .....	16

## Abbreviations

ABS	Australian Bureau of Statistics
AIHW	Australian Institute of Health and Welfare
AJCC	American Joint Committee on Cancer
COB	Country Of Birth
DOH	Department of Health Western Australia
FNA	Fine Needle Aspiration
ICD-O	International Classification of Disease- Oncology
ICD-10	International Classification of Disease 10 <sup>th</sup> Revision
MRI	Magnetic Resonance Imaging
NOS	Not Otherwise Specified
WA	Western Australia
WACR	Western Australian Cancer Registry

## Purpose

The purpose of this: Synthetic WA Cancer Registry Data Dictionary is to detail the synthetic data elements derived from the WA Cancer Registry (WACR).

## WACR Background

Since 1982, the Western Australian Cancer Registry has collected population-based incidence and mortality cancer data for use in the planning of health care services and the support of cancer monitoring, evaluation and research at local, national and international levels.

Examining specialists who identify the presence of malignancy in specimens are legally compelled to notify the WA Cancer Registry who will register the case, collecting a suite of demographic and clinical information according to national and international collection rules.

## Generation of WACR synthetic data

The generation of WACR Synthetic data at the Department of Health is performed in accordance with the data definitions.

The scope of the synthetic data is stomach cancer cases, diagnosed in WA, from 2010 to 2020. The size of the synthetic data is 2,228 records.

## Data definitions

This section provides specific information about synthetic data elements derived from the WACR, including definitions, permitted values, guide for use, and operational information.

All information relating to data elements in this data dictionary is specific to the WACR, and caution should be taken if these data elements are compared with those of other data collections. Where relevant, related national definitions have been referenced.

## Aboriginal Status

<b>Field Name:</b>	aboriginal_status
<b>Definition:</b>	Whether a person identifies as being of Aboriginal and/or Torres Strait Islander origin.
<b>Data Type:</b>	N
<b>Format:</b>	N(1)
<b>Permitted Values:</b>	0 Not Aboriginal and/or Torres Strait Islander 1 Aboriginal and/or Torres Strait Islander

### Guide for use

Aboriginal Status is provided on most pathology reports and may be updated from hospital notifications.

This metadata item is based on the Australian Bureau of Statistics (ABS) standard for Indigenous status.

In the synthetic dataset, the values "Aboriginal", "Torres Strait Islander", "Indigenous NOS" and "Aboriginal and Torres Strait Islander" have been combined to represent "Aboriginal and/or Torres Strait Islander".

"Non-indigenous", "Unknown / not stated" and NULL values have been combined to represent "Not Aboriginal and/or Torres Strait Islander"

### Examples

	Aboriginal Status
A person who identifies as Aboriginal has a specimen collected and sent to a pathology lab for investigation.	<b>1 Aboriginal Torres Strait Islander origin</b>
A person has a pathology specimen collected and sent to a lab for investigation but does not indicate their Aboriginal status.	<b>0 Not Aboriginal and/or Torres Strait Islander</b>

## Age Range

<b>Field Name:</b>	age
<b>Definition:</b>	Patient age at the time of tumour diagnosis.
<b>Data Type:</b>	Numeric
<b>Format:</b>	X (100)
<b>Permitted Values:</b>	N/A

### Guide for use

In the synthetic data ages are grouped in 5-year brackets and displayed as the lowest number in the bracket.

Ages between 0-4 Age in data set = 0

Ages between 5-9 Age in data set = 5

Ages between 10-14 Age in data set = 10

Ages between 15-19 Age in data set = 15

Ages between 20-24 Age in data set = 20

Ages between 25-30 Age in data set = 25

Continuing on for all ages

## Age Range at death

<b>Field Name:</b>	age_at_death
<b>Definition:</b>	The age of the person in (completed) years at the date of their death.
<b>Data Type:</b>	N
<b>Format:</b>	nvarchar(5), null
<b>Permitted Values:</b>	N/A

### Guide for use

In the synthetic data ages are grouped in 5-year brackets and displayed as the lowest number in the bracket

Ages between 0-4 Age in data set = 0

Ages between 5-9 Age in data set = 5

Ages between 10-14 Age in data set = 10

Ages between 15-19 Age in data set = 15

Ages between 20-24 Age in data set = 20

Ages between 25-30 Age in data set = 25

Continuing on for all ages

## Cancer type

<b>Field Name:</b>	cancer_type
<b>Definition:</b>	Grouping used for reporting purposes of cancer type based on cancer site (e.g., for breast, lung, prostate), morphology (e.g., lymphoma, leukemia), and behaviour.
<b>Data Type:</b>	String
<b>Format:</b>	X(1)
<b>Permitted Values:</b>	C16 Stomach

## Guide for use

Cancer type is aggregated from the combination of tumour site, morphology and behaviour. Most of these aggregations align with national and international aggregates, however some subgroups for haematopoietic cancers are unique to the Western Australian Cancer Registry.

In the synthetic data only cancer type C16 has been synthesised.

## Cause of death - cancer related

<b>Field Name:</b>	cause_of_death_cancer_related
<b>Definition:</b>	The disease or injury which were the underlying condition that led directly to death.
<b>Data Type:</b>	N
<b>Format:</b>	nvarchar(1), null
<b>Permitted Values:</b>	C - Died due to cancer X - Died due to cause other than cancer NULL - not collected

## Guide for use

This field indicates whether the underlying cause of a person's death was due to a cancer or not.

The WACR is notified about a person's death by forensic pathologists, the WA Registrar General (RG), through data linkage, and/or via the National Death Index (NDI). The RG may provide an initial death certificate stating that the cause of death is pending and follows it up with a subsequent death certificate with a definitive cause of death. Clinical coders update cause of death as more recent information is received.

Note that the registry's cause of death may not align with ABS cause of death because the registry often has access to a richer breadth of clinical information to ascertain the underlying cause of death.

## Cause of death - due to tumour

<b>Field Name:</b>	cause_of_death_due_to_tumour
<b>Definition:</b>	The underlying cause of death as indicated on the death certificate in accordance with ICD-10 Rules and guidelines for mortality and morbidity coding.
<b>Data Type:</b>	N
<b>Format:</b>	bit, null
<b>Permitted Values:</b>	1 - Yes 0 – No NULL - not collected

### Guide for use

This field is used to identify the cancer that is the underlying cause of death in accordance with ICD-10 Rules and guidelines for mortality and morbidity coding.as. Unlike the other cause of death fields that provide the death code and description, this field is a Yes/No response.

Note that one person may have multiple primary cancers over their lifetime and their cause of death should be attributed to one tumour, however there are some historical cases where this was ambiguous, and more than one tumour was flagged as resulting in their death. In this case, the earliest, or first tumour should be selected as resulting in the persons death.

The underlying cause of death is central to the registry's mortality coding and is comparable to international mortality reporting. • Code 1 – Yes indicates that the cause of death was due to a tumour, • Code 0 – No indicates that cause of death was not due to a tumour, and • NULL indicates the information was not collected or is missing.

### Detection method

<b>Field Name:</b>	detection_method
<b>Definition:</b>	Method of detection for the cancer registered
<b>Data Type:</b>	N
<b>Format:</b>	nvarchar(5), null
<b>Permitted Values:</b>	1 - Clinical, N.O.S 2 - Clinical, symptomatic 3 - Clinical, incidental 4 - Autopsy, N.O.S 7 - Screening 9 - Unknown

## Guide for use

This flag identifies the method of detection used to identify the cancer at diagnosis in question. Note that WACR does not have access to all these data sources so, for example, screening detected cancers would not be complete for this field.

## Diagnosis Postcode Region

<b>Field Name:</b>	diagnosis_postcode_region
<b>Definition:</b>	Western Australian region of residence at time of diagnosis.
<b>Data Type:</b>	String
<b>Format:</b>	X(4)
<b>Permitted Values:</b>	Metro Perth Metropolitan Region South Southern Region of WA North Northern Region of WA Postcodes not covered by the above Other categories (unknown or interstate)

## Guide for use

In the synthetic data, diagnosis\_postcode\_region is derived from diagnosis\_postcode that exists in the real data. It is reported by the patient as their residence at the time the tumour is diagnosed. In the synthetic data diagnosis\_postcode was aggregated to regions of Western Australia using the following mapping:

- Metro: (Perth Metropolitan Region): Postcodes 6000–6038, 6050–6084, 6090–6182, and 6556–6558.
- South: (Southern Region of WA): Postcodes 6200–6710.
- North: (Northern Region of WA): Postcodes 6041–6044 and 6713–6770.
- Other: Postcodes not covered by the above categories, unknown, or interstate.

## Grade

<b>Field Name:</b>	grade
<b>Definition:</b>	A description of a tumour based on how abnormal the cancer cells and tissue look under a microscope and how quickly the cancer cells are likely to grow and spread.
<b>Data Type:</b>	String
<b>Format:</b>	nvarchar(5), null
<b>Permitted Values:</b>	1 - Low / Well differentiated 2 - Int. / Moderately differentiated 3 - High / Poorly differentiated 4 - Anaplastic / Undifferentiated 9 - Unknown X - Not applicable NULL - not collected

## Guide for use

Tumour grade is an indicator of how quickly a tumour is likely to grow and spread. If the cells of the tumour and the organization of the tumour's tissue are close to those of normal cells and tissue, the tumour is called "well-differentiated." These tumours tend to grow and spread at a slower rate than tumours that are "undifferentiated" or "poorly differentiated," which have abnormal-looking cells and may lack normal tissue structures. Based on these characteristics and other differences in microscopic appearance, doctors assign a numerical "grade" to most cancers. However, there are some exceptions, cancers such as lymphoma, melanoma etc. don't have a grade, while Brain tumours grading is captured in the morphology. The factors used to determine tumour grade can vary between diverse types of cancer. Tumour grade is different from the stage of a cancer. Cancer stage refers to the size and/or extent (reach) of the original (primary) tumour and whether cancer cells have spread in the body. Cancer stage is based on factors such as the location of the primary tumour, tumour size, regional lymph node involvement (the spread of cancer to nearby lymph nodes), and the number of tumours present.

## Laterality

<b>Field Name:</b>	laterality
<b>Definition:</b>	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.
<b>Data Type:</b>	String
<b>Format:</b>	nvarchar(1), null
<b>Permitted Values:</b>	N - Not applicable

## Guide for use

This information is collected for the purpose of differentiating the site of the primary cancer. For example, a woman may present with a primary cancer in the left breast. She may return at a later stage with a new primary cancer in the right breast.

## Person Identifier

<b>Field Name:</b>	Person_ID
<b>Definition:</b>	This ID can be used to represent a single person in the WACR melanoma dataset. This is a unique sequential ID.
<b>Data Type:</b>	Number
<b>Format:</b>	N
<b>Permitted Values:</b>	6000 to NNNNNN

## Recurrence - date of onset

<b>Field Name:</b>	recurrence_date_of_onset
<b>Definition:</b>	Definition: Recurrence flag for date of onset.
<b>Data Type:</b>	N
<b>Format:</b>	date, null
<b>Permitted Values:</b>	yyyy-mm-dd

## Guide for use

The date that recurrence of an invasive cancer was identified. Whilst recurrence is reported where documentation from pathology or clinical information sources indicate recurrence, this information is not routinely provided to WACR and the presence of this field should not be used as an estimate of recurrence in the population.

## Region of Birth

<b>Field Name:</b>	region_of_birth
<b>Definition:</b>	Region of birth is the name of the region where the patient was born. ABS standard minor regions.
<b>Data Type:</b>	X
<b>Format:</b>	X (4)
<b>Permitted Values:</b>	As per ABS standard

## Guide for use

Country of birth is collected from secondary sources including the hospital morbidity data system (hospitalisations). Country of birth has been aggregated to ABS minor regions to create the region of birth. This field has been supplied to provide some measure of cultural and linguistically diverse patients.

## Sex

<b>Field Name:</b>	sex
<b>Definition:</b>	Description associated with patient's collected sex.
<b>Data Type:</b>	N
<b>Format:</b>	N(1)
<b>Permitted Values:</b>	1 Male 2 Female

## Guide for use

This is the sex as reported on the initial notification provided to the registry. Missing fields may be updated from hospitalisation data sources.

## Tumour Site - Code

<b>Field Name:</b>	tumour_site_code
<b>Definition:</b>	ICD-O code assigned to the stomach tumour topography or anatomical location on the body
<b>Data Type:</b>	N
<b>Format:</b>	(N4)
<b>Permitted Values:</b>	C160 C161 C162 C163 C164 C165 C166 C168 C169

## Guide for use

Tumour site refers to the anatomical part of the body where the stomach tumour is located. Each part of the body is assigned a specific code

The data present is coded according to ICD-O Version 3.2

Codes contain updates agreed and applied as per National Australian Association of Cancer Registries meetings.

## Year of Diagnosis

<b>Field Name:</b>	diagnosis_year
<b>Definition:</b>	The year the tumour was diagnosed
<b>Data Type:</b>	N
<b>Format:</b>	NNNN
<b>Permitted Values:</b>	2010 to 2022

### Guide for use

The earliest date associated with the tumour diagnosis is determined by the earliest health information notification received by the registry which confirms the presence of a malignancy. Typically this will be determined from the pathology reports.

This data set contains tumours diagnosed in the years from and including 2010 to 2022.

Produced by:  
Information and Performance Governance  
Information and System Performance Directorate  
Purchasing and System Performance Division  
The Department of Health Western Australia

**This document can be made available in alternative formats on request for a person with disability**

© Department of Health, State of Western Australia (2025).

Copyright to this material is vested in the State of Western Australia unless otherwise indicated. Apart from any fair dealing for the purposes of private study, research, criticism or review, as permitted under the provisions of the *Copyright Act 1968*, no part may be reproduced or re-used for any purposes whatsoever without written permission of the State of Western Australia.