Cancer Council Victoria

Victorian Cancer Registry

Guidelines for submitting data for record linkage

April 2020

Purpose

Described below is the preferred data format, for the data to be transferred to the VCR for record linkage purposes. Meeting the format as closely as possible will assist the VCR to more efficiently process your request and reduce processing costs.

1.   Data format

The preferred file format for submission of data for linkage is a tab delimited text file. If this is not possible then a .csv file is acceptable.

The data submitted must contain unique person records only i.e. one record per person.

Note: The tab delimited file must not contain any carriage returns or line feed characters within data strings e.g. char(10), char(13), char(20) or similar characters.

2.   Data specifications

The minimum data set for cases submitted to VCR for linkage includes a unique person records with cohort ID, full names, gender, date of birth and address. Additional elements such as UR number and hospital, other names, previous address and/or diagnosis details will increase the chance of matching to VCR and assist with clerical review to confirm or reject possible matches.

Field Name	Format	Comments/ required format
IDENTIFIERS		
StudyID		The unique record ID from the source data file Note, this identifier is mandatory and is used as a reference for seeking additional information from the data provider
URN		Where the source data unique identifier is not a hospital unit record number (URN), and a hospital URN has been collected, please provide this identifier
URN_Source		When supplying a hospital UR number for multiple sites, please include the name of the site for each UR number supplied
PERSONAL DETAILS		
Surname	Upper case	A-Z, no spaces, hyphens or apostrophes
First_Name	Upper case	First given name; A-Z, no spaces, hyphens, quotes or apostrophes
Second_Name	Upper case	Second given name; A-Z, no spaces, hyphens, quotes or apostrophes
Third_Name	Upper case	Other given names

PERSONAL DETAILS continued		
BirthDate	dd/mm/yyyy	E.g. 01/01/1980 Two-digit day of birth with leading zeros. Two-digit month of birth with leading zeros. Use 01 for unknown
Approx_BirthDate	0=N, 1=Y	Flag to indicate an approx. birth date has been supplied
Sex	1, 2, 3, 9	1 = Male; 2 = Female; 3 = Indeterminate; 9 = Unknown
Country of Birth		Country of birth code. Please provide a reference file of codes and descriptions
ADDRESS DETAILS		
PropertyName	Upper case	Property/Building name: no quotes / apostrophes (" ")
Flat_Number	Upper case	Flat/Unit number e.g. Unit 2 (Unit 2/7 North Rd)
Street number	Upper case	Street number e.g. 7 (Unit 2/7 North Rd)
Part number	Upper case	Street part e.g. A (63A Jones Ave)
Street name	Upper case	Street name e.g. JONES (63A Jones Ave)
Street Type	Upper case	ST, AVE, RD etc.
Suburb	Upper case	
Postcode		
State	Upper case	Applicable when providing data from multiple states
CANCER DETAILS		
Date_Diagnosis	dd/mm/yyyy	See BirthDate for format
Approx_DateDiagnosis	0 or 1	Flag to indicate an approximate date of diagnosis has been supplied (0 = No, 1 = Yes)
Cancer_Site		Please supply any text description or ICD codes that describe the cancer diagnosis. If supplying site in coded format, please indicate which ICD version has been used.
Cancer_Morph		Morphology of the tumour e.g. 8140/3
VITAL STATUS		
Date_Last_Contact	dd/mm/yyyy	Date that person was last known to be alive e.g. 01/01/2003
Vital_Status	A or D or U	A = Alive; D = Deceased; U = Unknown
DateDeath	dd/mm/yyyy	See BirthDate for format
Approx_DateDeath	0 or 1	Flag to indicate this is an approximate date of death has been supplied (0 = No, 1 = Yes)
Cause_Death		Where available, please supply a text description for the cause of death or code
ALTERNATE NAMES		
Alias_Surname		Other surnames that this person is known by including Maiden name; no spaces, hyphens, quotes or apostrophes
Alias_Given_Names		Other given names that this person is known by; no spaces, hyphens, quotes or apostrophes
Previous_Address		Address history if known; no quotes / apostrophes

[1] An alternative format for address is to supply an Address_Line by combining all numbers with the street name and street type. Please supply suburb and postcode as separate fields.

3. Data submission

All data submitted to and released by VCR is via the Registry's secure web portal. Please contact VCR (vcr@cancervic.org.au) to obtain access to this site and to arrange submission of your data for linkage.