

# Statistical linkage key 581 cluster

## Identifying and definitional attributes

**Metadata item type:** Data Set Specification

**METEOR identifier:** 750410

**Registration status:**

- [Health](#), Standard 17/12/2021

**DSS type:** Data Element Cluster

**Scope:** Statistical data linkage refers to the bringing together of data from different sources to gain a greater understanding of a situation or individual from the combined (or linked) dataset. This facilitates a better understanding of the patterns of service use by groups of clients for research, statistical or policy analysis, planning and evaluation purposes.

Its form is: **XXXXXDDMMYYYYX**

The sequence in which the linkage key is completed is as follows:

- Family name (the first 3 Xs)
- Given name (the 4th and 5th X)
- Date of birth by day (DD), month (MM) and 4-digit year (YYYY)
- Sex (last X).

### **XXX 2nd, 3rd and 5th letters of the family name.**

In the first three spaces the agency should record the 2nd, 3rd and 5th letters of the person's family name.

For example: If the person's family name is Smith the reported value should be MIH. If the person's family name is Jones the reported value should be ONS.

Regardless of the length of a person's name, the reported value should always be three characters long. If the legal family name is not long enough to supply the requested letters (i.e. a legal family name of less than five letters) then agencies should substitute the number '2' to reflect the missing letters. The placement of a number '2' should always correspond to the same space that the missing letter would have within the 3-digit field. A number (rather than a letter) is used for such a substitution in order to clearly indicate that an appropriate corresponding letter from the person's name is not available.

Cases where the family name has less than 5 letters:

If a person's family name is Farr, then value reported would be AR2 because the 2 is substituting for a missing 5th letter of the family name. Similarly, if the person's family name was Hua, then the value reported would be UA2 because the 2 is substituting for the missing 5th letter of the family name.

If a person's family name is missing altogether the agency should record the number 999 for all three spaces associated with the family name, (not the number 2). In some cultures, it is traditional to state the family name first. To overcome discrepancies in recording/reporting that may arise as a result of this practice, agencies should always ask the person to specify their legal first given name and their legal family name separately. These should then be recorded as first given name and family name as appropriate, regardless of the order in which they may be traditionally given.

If the person's family name includes non-alphabetic characters—for example

hyphens (as in Lee-Archer), apostrophes (as in O'Mara) or blank spaces (as in De Vries)—these non-alphabetic characters should be ignored when counting the position of each character.

### **XX 2nd and 3rd letters of the given name**

In the fourth and fifth spaces the agency should record the 2nd and 3rd letters of the person's given name.

For example: If the person's given name is Elizabeth the reported value should be LI. If the person's given name is Robert, the reported value should be OB.

If the person's given name includes non-alphabetic characters—for example hyphens (as in Jo-Anne) or apostrophes (as in D'Arcy), these non-alphabetic characters should be ignored when counting the position of each character.

Regardless of the length of a person's given name, the reported value should always be two characters long. If the given name of the person is not long enough to supply the requested letters (i.e. a name of less than three letters) then agencies should substitute the number '2' to reflect the missing letters. The placement of a number '2' should always correspond to the same space that the missing letter would have within the 2-digit field. A number (rather than a letter) is used for such substitutions in order to clearly indicate that an appropriate corresponding letter from the person's name is not available.

For example: If the person's legal name was Jo then the value reported would be O2 because the 2 is substituting for the missing 3rd letter of the given name.

If the person's given name is missing altogether the agency should record 99 for the two spaces associated with the given name. In some cultures it is traditional to state the family name first. To overcome discrepancies in recording/reporting that may arise as a result of this practice, agencies should always ask the person to specify their given name and their family name separately. These should then be recorded as first given name and family name as appropriate, regardless of the order in which they may be traditionally given.

### **Date of Birth**

In the sixth to thirteenth spaces the agency should record the day, month and year a person was born as follows:

- DD represents the day in the month a person was born
- MM represents the month in the year a person was born
- YYYY represents the year a person was born.

If date of birth is not known or cannot be obtained, provision should be made to collect or estimate age. Collected or estimated age would usually be in years for adults and to the nearest three months (or less) for children aged under 2. Additionally, an estimated date flag or a [date accuracy indicator](#) should be reported in conjunction with all estimated dates of birth.

For data collections concerned with children's services, it is suggested that the estimated date of birth of children aged under 2 should be reported to the nearest 3 month period, i.e. 0101, 0104, 0107, 0110 of the estimated year of birth. For example, a child who is thought to be aged 18 months in October of one year would have his/her estimated date of birth reported as 0104 of the previous year. Again, an estimated date flag or [date accuracy indicator](#) should be reported in conjunction with all estimated dates of birth.

### **Sex**

In the fourteenth space the agency should record the sex of the person. The term 'sex' refers to a person's biological characteristics. Codes 1 or M are used to record 'Male'. Codes 2 or F are used to record 'Female'. Codes 3 or X are used to record 'Another term'. Code 9 is used to record 'Not stated/inadequately described'.

Data that has been produced by linkage for statistical and research purposes should not be used subsequently for client management purposes.

This data cluster contains a set of specific data elements to be reported on in a

predetermined combination.

## Collection and usage attributes

**Guide for use:** It should be recognised that in some cases an individual may choose to report their gender when sex is being requested.

**Implementation start date:** 01/07/2018

## Source and reference attributes

**Submitting organisation:** Australian Institute of Health and Welfare.

**Steward:** [Australian Institute of Health and Welfare](#)

**Origin:** AIHW (Australian Institute of Health and Welfare) 1998. Home and Community Care (HACC) Data Dictionary Version 1.0. Report prepared for the Commonwealth and State/Territory government HACC Officials.

**Reference documents:** National Community Services Information Management Group 2004. Statistical data linkage in community services data collections. Canberra: AIHW.

## Relational attributes

**Related metadata references:** Supersedes [Statistical linkage key 581 cluster](#)

- [ACT Health \(retired\)](#), Candidate 09/08/2018
- [Health](#), Superseded 17/12/2021
- [Homelessness](#), Standard 10/08/2018

See also [Date—accuracy indicator, code AAA](#)

- [ACT Health \(retired\)](#), Candidate 09/08/2018
- [Community Services \(retired\)](#), Standard 30/09/2005
- [Disability](#), Standard 07/10/2014
- [Early Childhood](#), Standard 21/05/2010
- [Health](#), Standard 04/05/2005
- [Homelessness](#), Standard 23/08/2010
- [Housing assistance](#), Standard 23/08/2010

**Implementation in Data Set Specifications:**

[Alcohol and other drug treatment services NMDS 2022–23 Health](#), Standard 17/12/2021

*Implementation start date:* 01/07/2022

*Implementation end date:* 30/06/2023

[Alcohol and other drug treatment services NMDS 2023-24 Health](#), Qualified 31/08/2022

*Implementation start date:* 01/07/2023

*Implementation end date:* 30/06/2024

## Metadata items in this Data Set Specification

Seq No.	Metadata item	Obligation	Max occurs
1	<a href="#">Person—letters of family name, text XXX</a>	Mandatory	1
2	<a href="#">Person—letters of given name, text XX</a>	Mandatory	1
3	<a href="#">Person—date of birth, DDMMYYYY</a>	Mandatory	1
4	<a href="#">Person—sex, code X</a>	Mandatory	1
5	<a href="#">Record—linkage key, code 581 XXXXXDDMMYYYYX</a>	Mandatory	1
6	<a href="#">Date—accuracy indicator, code AAA</a>	Conditional	1

© Australian Institute of Health and Welfare 2015–2022

This product, excluding the AIHW logo, Commonwealth Coat of Arms and any material owned by a third party or protected by a trademark, has been released under a Creative Commons BY 3.0 (CC BY 3.0) licence. Excluded material owned by third parties may include, for example, design and layout, images obtained under licence from third parties and signatures. We have made all reasonable efforts to identify and label material owned by third parties.

You may distribute, remix and build upon this work. However, you must attribute the AIHW as the copyright holder of the work in compliance with our attribution policy available at [www.aihw.gov.au/copyright](http://www.aihw.gov.au/copyright). The full terms and conditions of this licence are available at <http://creativecommons.org/licenses/by3.0/au/>.

Enquiries relating to copyright should be addressed to the Head of the Communications, Media and Marketing Unit, Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601.