



Australian Government
**Australian Institute of
Health and Welfare**

Pathways in Aged Care 2014: technical guide

The logo for the Australian Institute of Health and Welfare (AIHW), consisting of the letters 'AIHW' in a bold, sans-serif font. Each letter is a different color: 'A' is teal, 'I' is green, 'H' is blue, and 'W' is purple.

AIHW



Australian Government

**Australian Institute of
Health and Welfare**

Pathways in Aged Care 2014

Technical guide

Australian Institute of Health and Welfare
Canberra

Cat. no. AGE 82

The Australian Institute of Health and Welfare is a major national agency whose purpose is to create authoritative and accessible information and statistics that inform decisions and improve the health and welfare of all Australians.

© Australian Institute of Health and Welfare 2017



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/>. The full terms and conditions of this licence are available at <<http://creativecommons.org/licenses/by/3.0/au/>>.

A complete list of the Institute's publications is available from the Institute's website <www.aihw.gov.au>.

Suggested citation

Australian Institute of Health and Welfare 2017. Pathways in aged care 2014: technical guide. Cat. no. AGE 82. Canberra: AIHW.

Australian Institute of Health and Welfare

Board Chair
Mrs Louise Markus

Director
Mr Barry Sandison

Any enquiries relating to copyright or comments on this publication should be directed to:

Australian Institute of Health and Welfare
GPO Box 570
Canberra ACT 2601
Tel: (02) 6244 1000
Email: info@aihw.gov.au

Published by the Australian Institute of Health and Welfare

This publication is printed in accordance with ISO 14001 (Environmental Management Systems) and ISO 9001 (Quality Management Systems). The paper is sourced from sustainably managed certified forests.



**Please note that there is the potential for minor revisions of data in this report.
Please check the online version at <www.aihw.gov.au> for any amendments.**

Contents

- Acknowledgments.....iv**
- Abbreviationsv**
- 1 Introduction1**
- 2 Developing the PIAC link map.....2**
 - 2.1 Included aged care programs2
 - 2.2 Data sources for PIAC 20144
- 3 Overview of linkage.....9**
 - 3.1 Linkage strategy9
 - 3.2 Client identification and data for linkage.....10
 - 3.3 Quality of the data available for linkage13
 - 3.4 Linkage quality.....13
- 4 Program and data issues affecting analysis15**
 - 4.1 Residential aged care15
 - 4.2 Transition Care Programme (TCP)17
 - 4.3 Home care packages.....17
 - 4.4 Program leave18
 - 4.5 Aged Care Assessment Program (ACAP).....18
 - 4.6 Home and Community Care (HACC)19
- 5 PIAC special analysis data sets22**
 - 5.1 Demographic characteristics.....22
 - 5.2 Service use dates23
 - 5.3 ACAP assessments with cleaned dates.....25
 - 5.4 Care pathways.....26
- 7 Accessing PIAC data and information30**
- Appendix A: Data Quality Statements31**
- References.....32**
- List of tables33**
- Related publications34**

Acknowledgments

This technical user guide was written by Rosemary Karmel and is an update to the Pathways in Aged Care (PIAC) user guide published in 2016, *Introduction to Pathways in Aged Care 2014* and includes a number of the methods sections from *Patterns in use of aged care: 2002–03 to 2010–11* (AIHW 2014).

Three members of the Data Linkage Unit at the Australian Institute of Health and Welfare (AIHW) have been responsible for developing the PIAC link map, both the earlier version PIAC 2011 and the current PIAC 2014 update: Phil Anderson guided the project throughout; Rosemary Karmel was responsible for the day to day project management, and also prepared much of the data for linkage; and Andrew Powierski undertook the data linkage. In addition, Cath Lawrence and Richard Solon of the Data Linkage Unit assisted with preparation of data for linkage to be used for PIAC 2011.

Before the data linkage was undertaken, ethics approval and permission to use the required aged care data were obtained from the relevant Australian and state and territory government departments. This project also linked data to the National Death Index to identify deaths. The registries of births, deaths and marriages and the National Coronial Information System are acknowledged as the providers of these data and the AIHW thanks them for enabling access.

To protect the privacy of individuals, the linkage was carried out within the AIHW using the Institute's data linkage protocol for linking data sets held within the AIHW.

Over the years, people in the Ageing and Aged Care Unit at AIHW (now incorporated into the Disability and Ageing Unit) have provided invaluable advice on aged care programs and on the use of the related program data. In particular, former staff members Evon Bowler and Peter Braun are thanked for their advice and assistance in the early years.

Abbreviations

ACAP	Aged Care Assessment Program
ACAT	Aged Care Assessment Team (for ACAP)
ACCR	Aged Care Client Record
ACFI	Aged Care Funding Instrument
ADL	activities of daily living
AIHW	Australian Institute of Health and Welfare
BEH	behavioural
CACP	Community Aged Care Package
CDC	Consumer Directed Care
CHC	complex health care
DOB	date of birth
DOD	date of death
EACH	Extended Aged Care at Home
EACHD	Extended Aged Care at Home Dementia
FMR	false match rate (estimated)
HACC	Home and Community Care
HCP	Home Care Packages Program
KBL	key-based linkage
m_tf	estimated marginal trade-off between additional true and additional false matches for links established using the match key when compared with matches made by a slightly more precise key
max_FMR	maximum FMR for determining number of versions that can be compared in KBL
MDS	minimum data set
NACDC	National Aged Car Data Clearinghouse
NACDC14	2014 provision of the NACDC to AIHW
NACDC15	2015 provision of the NACDC to AIHW
NBL	name-based linkage
NDI	National Death Index
PIAC	Pathways in Aged Care
PRAC	permanent residential aged care
p_id	PIAC person identifier
PPV	positive predictive value
RAC	residential aged care
RCS	Resident Classification Scale
RRAC	respite residential aged care

SLK	statistical linkage key
SLK-581	statistical linkage key, comprising 5 letters of name, 8 digits of date of birth and sex
SPARC	an abbreviation covers RAC, EACH, EACHD and TCP clients as the client identifiers for these people all begin with the letters 'SPARC' on the NACDC.
TCP	Transition Care Programme
VHC	Veterans' Home Care

1 Introduction

Over the last 30 years, policy direction, and consequently program development, has been influenced by a small number of underlying principles. These include that many older Australians: prefer to live in the community rather than in residential care; prefer to 'age in place' rather than change residence when care needs change; and want aged care services to be flexible and accessible. Since the early 1980s the Australian Government has implemented a range of reforms that have increasingly moved the focus of care provision from residential aged care to include a wide range of community-based aged care services. Exploring how Australians access these programs and move between them is an important part of understanding the complex needs of older Australians and how the aged care system is meeting them.

While national unit record level data have been available for most of the aged care programs since July 2006, the data collections for the different programs do not use a common client identifier, and so are not fully integrated. Consequently, most analyses are program-specific. To overcome the limitations arising from this lack of integration, key community care, residential care, assessment program and deaths data have been brought together using statistical data linkage processes. Data from aged care service programs, assessments for the use of aged care and the National Death Index were linked to create the Pathways in Aged Care (PIAC) link map. This database is suitable for person-based analysis of aged care pathways and patterns of program use over time.

PIAC originally covered aged care assessments and use of 7 aged care service programs, as well as deaths from 1 July 2002 to 30 June 2011. The programs included:

- Aged Care Assessment Program (ACAP)
- Permanent and respite residential aged care (PRAC and RRAC)
- Home and Community Care (HACC)
- Veterans' Home Care (VHC)—not included in PIAC 2014
- Home-based aged care packages: Community Aged Care Packages (CACP), Extended Aged Care at Home (EACH) and Extended Aged Care at Home Dementia (EACHD)
- Transition Care Programme (TCP)
- National Death Index (NDI).

PIAC 2014 is an expansion of the link map to include all program use reported in linkable data sets and deaths between 1 July 1997 and 30 June 2014. In particular, it includes data for the Home Care Packages (HCP) Program which replaced CACP, EACH and EACHD on 1 August 2013 (collectively, these 4 are now referred to as 'home care packages'). However, PIAC 2014 does not include expanded VHC data.

This document is designed to support users of PIAC and provides an introduction to the linkage processes and a description of the various aged care programs and data sources. Notes on data quality affecting use and exclusions for the aged care data and deaths data are also included.

2 Developing the PIAC link map

In order to facilitate investigations into how people access and move between aged care programs, as well as people's aged care pathways from first use until death, data on the use of key aged care service programs, Aged Care Assessment Teams (ACAT) assessments and deaths have been linked using statistical linkage processes to allow data from different programs to be combined for statistical analysis—the resulting database is termed the PIAC link map.

The current link map PIAC 2014 has built on, and benefited from, two earlier linkage projects:

- The PIAC cohort study: An initial link map which included program use between July 2003 and June 2006 for a cohort of 2003–04 ACAP clients undertaken by a consortium of researchers at the Australian Institute of Health and Welfare (AIHW), University of Queensland and La Trobe University, (AIHW 2009, 2010, 2011a, 2011b; Karmel et al. 2010, 2012).
- PIAC 2011: A second linkage project extended the link map derived for the PIAC cohort study to cover all program use and deaths during the period 1 July 2002 to 30 June 2011 (AIHW 2014, 2015).

The link map is in effect a set of concordance tables between the PIAC person identifier (p_id) and person and/or record identifiers used in the National Aged Care Data Clearinghouse (NACDC) tables, the pre-NACDC HACC tables, and the NDI.

PIAC 2014 has extended the linkage to include any aged care program use reported in linkable data sets between 1 July 1997 and 30 June 2014, and all deaths between 1 July 1997 and 30 September 2015 (to ensure deaths data are captured for people who had used aged care in the reference period). The aged care programs included in PIAC 2014 are the same as those included in earlier PIAC maps, with the exception of VHC, which is no longer included (Box 2.1).

2.1 Included aged care programs

Aged Care Assessment Program (ACAP)

The Aged Care Assessment Program has been operational since 1985. Under ACAP, multi-disciplinary ACATs determine people's care needs and make recommendations on preferred long-term living arrangements for clients. Relevant approvals are required from an ACAT in order to access funded places within many of the aged care programs and packages.

Residential aged care (RAC)

Residential aged care (Commonwealth-funded from 1963) provides both respite and permanent care in residential care facilities. An ACAT approval is required to access funded places.

Respite residential aged care (RRAC)

People who live in the community can receive short-term respite residential aged care in aged care facilities. Respite care is available on either a planned or emergency basis to older people who intend to return to their own home, but require temporary residential aged care.

It supports older people in transition stages of health, as well as providing carers with a break from their caring duties. Respite care is offered as either low- or high-care.

Permanent residential aged care (PRAC)

When people can no longer be supported to live independently in the community, they may move into permanent residential aged care. After a person enters permanent care, the resident is appraised to gauge their care needs, and hence their government subsidy level. From March 2008 these appraisals have been carried out using the Aged Care Funding Instrument (ACFI). This instrument measures the level of care required, calculating an overall score which is then used to determine a government subsidy amount. Residents may be re-appraised if their care needs change substantially.

From 1 July 2014, the distinction between low care and high care was removed for PRAC as part of the 2012 aged care reforms. To enter permanent residential aged care, individuals now only need an ACAT approval, and not for a particular level of care, and they no longer have to move between facilities as their care needs change from low to high.

Home and Community Care (HACC)

At the time, the largest government-funded aged care program was the Commonwealth HACC program, which provided a range of basic maintenance and support services for frail older people living in the community.

HACC has been operational since 1985 and brought together a number of separate programs operating from the mid-1950s under Commonwealth–state agreements. Previously funded jointly by the Commonwealth (Australian) and state and territory governments, on 1 July 2012 the Australian Government assumed full policy, funding, and day-to-day responsibility for HACC services for people aged 65 and over, and for Aboriginal and Torres Strait Islander people aged 50 and over in all states and territories except Victoria and Western Australia (termed 'Commonwealth HACC').

From July 2015, HACC became the main part of the Commonwealth Home Support Programme. An ACAT approval is not required for access.

Home care packages

Community Aged Care Packages (CACP)

The CACP program (operational from 1992, replaced by Home Care Packages Programme in 2013) provided support services for older people with complex needs living at home who were otherwise eligible for admission to 'low-level' residential care. The packages provided a range of home-based services, but not home nursing assistance and allied health services, with care being coordinated by the package provider. Access required an ACAT approval.

Extended Aged Care at Home (EACH)

The EACH program (operational from 2002, replaced by Home Care Packages Programme in 2013) provided care at home equivalent to 'high-level' residential care. Access required an ACAT approval.

Extended Aged Care at Home Dementia (EACHD)

The EACHD program (operational from 2006, replaced by Home Care Packages Program in 2013) provided a community care option specifically aimed at high-care clients with dementia and behavioural and psychological symptoms. Access required an ACAT approval.

Home Care Packages (HCP)

The Home Care Packages Program began on 1 August 2013, replacing the former packaged care programs of CACP, EACH, and EACHD. Four levels of packages are available, from Level 1, which supports people with basic care needs, to Level 4 which supports people with high-care needs. Home care packages are required to be delivered using Consumer Directed Care (CDC). CDC was phased in from 2013, with all home care packages using a CDC model of care from July 2015. As with the earlier package programs, an ACAT approval is required.

Transition Care Programme (TCP)

TCP (operational from 2005) provides short-term care to older people who are leaving hospital who are assessed as otherwise being eligible for at least low-level residential aged care. It aims to improve recipients' independence and functioning and delay entry into residential aged care. Access requires an ACAT approval. TCP care can be provided at home or in 'live-in' facilities, including residential aged care and hospital.

2.2 Data sources for PIAC 2014

The data sources for the various aged care programs and deaths are described on the next page. Issues related to scope, limitations and data quality are also discussed (Table 2.1).

Table 2.1: Overview of PIAC data sources

Scope of data	Data source	Notes
ACAP assessments		
<p>ACAP assessments that ended between 1 July 2003 and 30 June 2014</p>	<p>NACDC15</p> <p>NACDC14 included ACAP MDS data only up to 2012–13. However, data for ACAP 2013–14 was included in NACDC15, and so NACDC15 was used as the source for ACAP data.</p> <p>Note: NACDC15 was released after the name-based matching to SPARC/CACP and NDI was well underway, and so was not used for those.</p>	<p>Data on the provision of assessments carried out under ACAP are recorded in the ACAP Minimum Data Set (MDS) collection (from 1 July 2003). People are not assigned a unique administrative identifier, but a statistical linkage key (SLK) is recorded at each assessment to allow client-level statistical analysis of program use. ACAP MDS data are provided to the NACDC annually, but generally 12 months behind the other program data.</p> <p>2,280,580 ACAP MDS distinct assessments are linked to 1.16 million people in the PIAC 2014. Assessments with insufficient data for internal linkage were excluded from the linkage process, and consequently PIAC 2014.</p> <ul style="list-style-type: none"> • A record was considered to have insufficient data for internal linkage if it had more than 1 missing component of SLK-581 or any missing date of birth data. • Date of births recorded as 1 January 1900 and 1901 were treated as dummy values and were excluded from the linkage process. This mainly affects data from the earlier years. • A small number (8,058 which were included in PIAC 2011) of records on the original annual MDS data provisions for 2003–04 to 2010–11 could not be mapped to the data from the NACDC. These records are therefore not included in PIAC 2014. <p>Assessment level data for ACAP were collected under the ACAP MDS V2 from 1 July 2003 and 30 June 2014. However, data are incomplete in the early years. Specifically, implementation of ACAP MDS V2 data collection was completed during 2003–04 for all jurisdictions except New South Wales and Queensland. For these 2 states, implementation was completed by the end of 2005–06.</p> <p>Records with only 1 missing component of the statistical linkage key (SLK-581) were considered to have sufficient information for client identification if the record contained a postcode. Internal deterministic linkage was then used to see if these records should be combined with other records with complete data for SLK-581. If only sex were missing from the SLK-581, this internal matching process was carried out even if postcode were also missing. Records with missing or poor DOB were not considered to have sufficient information for client identification.</p> <p>An assessment can be recorded more than once on the ACAP MDS, usually due to updated information. The researcher will need to decide which version to use for analysis (commonly that with the latest effective_from_date).</p> <p>PIAC 2014 to ACAP MDS concordance only includes assessments reported on NACDC15. There were a number of ACAP assessments included in PIAC 2011 that were not on NACDC15: 97% of these people with assessments not on NACDC15 were from NSW (67%) or SA (30%), and 99% were for 2003–04.</p> <p>A number of ACAP MDS assessments (300,127 with 406,325 associated records) that were in NACDC15 have been excluded from the ACAP concordance for PIAC 2014. These were excluded for 3 main reasons: assessment end date was before 1 July 2003, assessment end date was after 30 June 2014, and poor name of DOB information for SLK-581.</p>

(continued)

Table 2.1 (continued): Overview of PIAC data sources

Scope of data	Data source	Notes
PRAC / RRAC / EACH / EACHD / TCP \ CACP		
People who used PRAC, RRAC, EACH, EACHD and TCP (recorded by source system SPARC) or CACP (recorded by source system CACP) between 1 July 1997 and 30 June 2014	National Aged Care Data Clearinghouse 2014 (NACDC14)	Person identifiers in the SPARC and CACP source systems are not integrated. However, these person identifiers are generally unique.
		<p>There are 62 cases where there was a SPARC ID on the PIAC 2011 master list but not on NACDC14, and 15 cases where there was a CACP ID on PIAC 2011 but not on NACDC14. These represent a small discrepancy between data provided via ACCMIS, as used for the original PIAC database, and data provided for NACDC14.</p> <p>Because people can use multiple programs, an individual may have both a SPARC and a CACP ID. There are 1.26 million people in PIAC 2014 who were either SPARC and/or CACP clients between 1 July 1997 and 30 June 2014.</p>
HCP		
People who used HCP between 1 August 2013 and 20 June 2014	National Aged Care Data Clearinghouse 2015 (NACDC15) as data were not available on NACDC14	People's use of HCP is recorded on separate NACDC tables using different client identifiers. These IDs are distinguished by the prefix ACMPS and are not directly integrated with SPARC or CACP IDs. However, the 'HCP' tables on NACDC also include all historical data for the pre-August 2013 programs.
		<p>People who migrated from the CACP, EACH and EACHD programs to HCP may be reported against both the HCP and earlier programs. The HCP client identifiers are not integrated with either those for SPARC or CACP clients, although there is a partial mapping table on NACDC15.</p> <p>Inspection of the HCP demographics table on NACDC15 suggests that individuals can get more than one HCP client ID.</p>

(continued)

Table 2.1 (continued): Overview of PIAC data sources

Scope of data	Data source	Notes
HACC		
<p>HACC services provided between 1 July 2001 and 30 June 2014 and that were reported on either the MDS V1.0 or 2.0 are included on PIAC 2014</p>	<p>For PIAC 2014, HACC MDS data for 2001–02 (V1.0) and for 2011–12 to 2013–14 (V2.0) were linked to PIAC 2011. As the NACDC only contains MDS V2.0 data, MDS V1.0 linkage data for 2001–02 came from the original provision of the annual data. Data for MDS V2.0 (i.e. for 2011–12 to 2013–14) come from the NACDC15—the most recent data provision at the time of linkage.</p>	<p>Data on the provision of services provided through HACC are recorded in the HACC MDS annual data collections (from 2001–02). Up to 2009–10, the HACC MDS was collected by the state and territory governments, and then collated into a national data set by the Australian Government.</p> <p>Clients are not assigned a unique administrative identifier, but an SLK (SLK-581) is reported on each service record to allow client-level statistical analysis of program use.</p> <p>Data on services provided through HACC were collected via the HACC MDS V1.0 for 2001–02 to 2004–05 and using HACC MDS V2.0 from 2005–06.</p>
		<p>Records with insufficient data for internal linkage were excluded from the linkage process, and consequently PIAC 2014.</p> <ul style="list-style-type: none"> • A record was considered to have insufficient data for internal linkage if it had more than 1 missing component of SLK-581 or any missing date of birth data. • Records with only carer SLK-851 data were not included in PIAC 2014. • As for PIAC 2011, 1 January 1900 and 1901 DOBs were treated as dummy values. Consequently, records with these dates of birth were excluded from the linkage process, and therefore PIAC 2014, as having insufficient data for internal linkage. This mainly affects data from the earlier years. • Other records with 1 January dates of birth that did not link to a record on the link map with full name data were also assumed to be dummy dates of birth and were not added to the map file. • A small number of records on the MDS V2.0 data provisions for 2005–06 to 2010–11 could not be mapped to the NACDC.

(continued)

Table 2.1 (continued): Overview of PIAC data sources

Scope of data	Data source	Notes
NDI (deaths)		
All deaths from 1 January 1997 to September 2015	<p>NDI as at September 2015</p> <p>The scope was made larger than the PIAC period to ensure that we identified, as far as possible, all deaths of PIAC clients.</p>	<p>The NDI is a database, housed at the AIHW, which contains records of all deaths occurring in Australia since 1980. The data are obtained from the Registrars of Births, Deaths and Marriages in each state and territory.</p> <p>The Index is designed to facilitate epidemiological studies. It contains name and demographic data to facilitate matching to other data sets, as well as a death registration number. The latter is used to derive a mortality ID which can be used to import causes of death from the National Mortality Database for projects with an appropriate Ethics Committee approval.</p> <p>2.56 million deaths are included in PIAC 2014. Exclusions include: those who were born and died on the same day according to the NDI, unreported deaths, and deaths reported after September 2015.</p>
		<p>Deaths may be reported more than once on the NDI. Multiple versions are generally assigned the same mortality ID. All versions have been retained to facilitate bringing in causes of death information (where approved). There are 2,559,167 distinct p_id–DOD pairs on the NDI concordance file for PIAC (1.0013 records per p_id with a death).</p> <p>Some people have more than 1 mortality ID. In a small number of cases, due to processing errors different people may have been given the same mortality ID. Where possible, these have been resolved. There are 2,558,201 distinct p_id–mortality ID pairs on the NDI concordance file for PIAC (1.0010 mortality ids per p_id with a death).</p> <p>In a small number of cases the date of death as reported on the NDI is not valid.</p>

3 Overview of linkage

Before data linkage was undertaken, ethics approval and permission to use the required data were obtained from all relevant bodies. In addition, to protect the privacy of individuals, the linkage was carried out within the AIHW using the Institute's data linkage protocol, see (AIHW 2006).

3.1 Linkage strategy

Data linkage is a powerful tool for identifying multiple appearances of individuals within a data set and for integrating client information across data sets. As the information recorded for an individual may vary from data set to data set—due to either differences in reporting (for example, in first name) or errors—a robust linkage process should allow for some discrepancy in characteristics. There are two main types of data linkage:

- Key-based record linkage, in which the linkage of records is based on exact agreement of the linkage variables. Variation in reporting can be allowed for by using a number of different keys.
- Probabilistic record linkage, in which the linkage of records in 2 files is based on the probabilities of agreement and disagreement between a range of linkage variables. Probabilistic linkage allows for variation in reporting by allowing probabilities of agreement to be less than 1 and probabilities of disagreement to be greater than 0.

Key-based linkage is commonly used when linking either using a person identifier or when full name data are not available but other data items are available which, when combined, can be used to link records. Probabilistic linkage is generally used when full name information, along with other demographic data, is available. Because only some of the data sets included in the PIAC database contain full name information, both types of linkage were used when developing the linked database.

Name-based linkage

Probabilistic name-based linkage (NBL) was used when linking data sets that both contained full name information. It involved running a series of passes that allow for variation in full name information and demographic data. Each pass consisted of deterministic pairwise matching on selected blocking variables and then calculating a weight based on probabilities of agreement and disagreement for the blocking and match variables for each respective match pair. Sample-based clerical review was then conducted across all passes to identify initial high and low weight cut-offs for matches where there was variation in reported match data (Guiver 2011). The final weight cut-off on which to base match decisions was obtained by further comparisons using reported postcode and date of death data (when available), again using sample-based clerical review. Finally, the links were examined for cross links (for example, one person on one data set matching to two or more on the other) and processed accordingly. A description of NBL process used is given Appendix B of *Patterns in use of aged care: 2002–03 to 2010–11* (AIHW 2014).

Key-based linkage

The ACAP and HACC data sets included in the PIAC project do not have full name information, but do contain the statistical linkage key SLK-581. Linkage of these data sets to the PIAC database was therefore undertaken using key-based linkage (KBL). The KBL

process for PIAC involves matching via multiple deterministic match passes, using linkage keys derived from data items available for linkage (that is, common to both data sets). Each pass uses a linkage key based on a different combination of the linkage variables. An algorithm is used to identify suitable linkage keys and the order in which they should be used. Using multiple passes with different linkage keys allows matches to be identified for units which have linkage variables reported differently in the two data sets. This method maximises the value of the SLK-581 for linkage and has been used in a number of projects (Karmel et al. 2010). Again, a step-by-step description of the match passes and clerical review processes used is Appendix B of *Patterns in use of aged care: 2002–03 to 2010–11* (AIHW 2014).

3.2 Client identification and data for linkage

Before undertaking data linkage, data sets with appropriate client identifiers and appropriate linkage variables had to be derived.

Client identification

Two processes were used to identify distinct clients within the data sets contributing to PIAC. The method used depended on whether the data set contained an administrative program client identifier.

Data sets with administrative person identifiers

In some aged care data sets, unique individuals can be identified relatively reliably. All but one of the data sets included in PIAC which have full name data (SPARC programs, CACP and NDI) also have a unique administrative person identifier (A_PID). However, even in these data sets it is possible for a person to have more than one identifier due to an administrative or processing error. Consequently, before linking, data sets with full name data were de-duplicated using the name-based linkage process by matching a data set to itself. A small number of people with more than one A_PID were identified in each data set. In these cases, for matching purposes the person was assigned a single A_PID.

For HCP, client ids with more than one ACAP assessment could have more than one client ID. As for the other named data sets, before linking, the HCP client list was de-duplicated using the name-based linkage process. A total of 6,301 duplicate pairs were identified. Again, for matching purposes the person was assigned a single A_PID.

Data used in name-based linkage includes not only names but also other demographic information (Box 3.1).

Box 3.1: Data used in name-based linkage:

Name-based linkage was used to link when full name data are available (that is, for SPARC programs, CACP, HCP and NDI). The data used in this process included:

- first name
- last name
- middle name
- other name
- date of birth
- sex
- (possible) date of death
- last seen date
- postcode of residence
- suburb of usual residence (used to obtain possible postcodes if postcode had not been reported).

Note that not all variables were available on all data sets.

A person's postcode used in linkage could change depending on the data sets being matched:

- When linking SPARC to the NDI, the preferred postcode for the SPARC person was that of the last known residence. For people permanently living in RAC this was the postcode of their RAC facility; for others it was that of their home address in the community.
- When linking to community care programs, the preferred postcode for the people being linked was that of the last known residence in the community. For people in PRAC, this was the postcode of their usual residence before they entered the facility.

As people can change where they live, both in the community and in residential care, a person can have several postcodes recorded in a data set. For example, the SPARC data can contain several postcodes relating to the same client over a year: the postcode of usual residence before going into PRAC and the postcodes of any facility the person used. This postcode variation was used when identifying matches among the 'possible links'. For example, when linking SPARC to NDI up to 3 postcodes were used: the client's postcode in the community before entering PRAC, the postcode of the last PRAC facility used and the postcode of where the client died according to the NDI.

Data sets without administrative person identifiers

Some aged care data sets do not readily identify individuals by name or other identifier. There is no unique program client identifier in either the ACAP or HACC MDS, and full name is not recorded. Rather, both collections contain data items through which repeat assessments by individuals can be identified with high probability; namely, SLK-581. Previous analysis has shown that SLK-581 distinguishes well between individuals in aged care data sets (AIHW: Karmel 2005a, 2005b; AIHW: Karmel & Braun 2004).

Although not common, different people can have the same SLK-581 (0.6% in a population of 440,000) (AIHW: Ryan et al. 1999:78). Therefore, as in PIAC 2011, to reduce the likelihood of combining data for different people—especially in the large HACC data sets—a client in

the annual ACAP and HACC data sets was defined by SLK-581 combined with the first digit of the client's postcode of usual residence, and a collection client identifier (C_CID) assigned accordingly. That is, essentially a client was defined by SLK-581 within a state or territory, except with New South Wales and the Australian Capital Territory being combined. Note that under this definition a person who moved during the year will be identified as two clients if, on moving, their postcode of usual residence changed in the first digit, that is, they would have two C_CIDs. If a client's postcode was missing, the client's state of usual residence was assumed to be the same as that of the service provider. The treatment of ACAP and HACC records with 1 or more missing component of SLK-581 is described in Table 2.1 above.

Data used in key-based linkage include letters of name and some key demographic and location information (Box 3.2).

Box 3.2: Data used in key-based linkage (KBL)

Key-based linkage was used to link ACAP and HACC annual data sets to the PIAC database. The data used in this process included:

- second, third and fifth letters of last name S235, and various paired combinations of these letters; that is S23, S25 and, S35
- second and third letters of first name F23
- day of birth (d)
- month of birth (m)
- year of birth (y)
- sex (s)
- full person postcode (pc4)
- first 3 digits of person postcode (pc3)
- first 2 digits of person postcode (pc2)
- first digit of person postcode (pc1)
- suburb (used to derive pc1–pc4 if postcode was not reported, or to derive alternative postcodes)
- date of last completed ACAP assessment in a financial year (when matching ACAP MDS)
- ID of ACAT undertaking the last completed assessment in a financial year (when matching ACAP MDS).

Note that not all variables were available for linkage on all datasets. In particular, ACAP assessment data were available only for SPARC and ACAP program data. In addition, to avoid false matches due to the large number of assessments undertaken during a year, ACAP assessment date was not used without ACAT ID.

Since ACAP provides approvals for entry into RAC or for the use of care packages, and HACC is a service for people living in the community, the preferred postcode of usual residence used in KBL was that which related to living in the community. For people in PRAC, this was the postcode of their usual residence before moving into a facility.

As people can move during the year, a person can have several postcodes recorded on the HACC or ACAP MDS. In such cases, all postcodes were included in the KBL process, with the priority of the postcode used in the KBL algorithm based on recency of use. In addition, for people in PRAC, the postcode of the facility was used as a (lower priority) alternative.

3.3 Quality of the data available for linkage

The presence of missing linkage data reduces the likelihood of identifying true matches. The number of missed matches will also be relatively high if there are unreliable data on one of the data sets. However, if both data sets being matched have similar processes for recording poor information (for example, recording dates of birth as 1 January of the year derived from current age) then the likelihood of making false matches decreases.

The quality of the linkage data was examined in detail for PIAC 2011, and is reported in Appendix B of *Patterns in use of aged care: 2002–03 to 2010–11* (AIHW 2014). In general, the data sets which included full name data were less likely to have missing name or date of birth information than those which contained the data for SLK-581 but not full name (ACAP and HACC MDSs). Under 0.2% of clients on data sets with full name data were missing either name or date of birth (DOB) information. From 2006–07, the ACAP MDS had similarly low numbers of records with insufficient data for linkage. In the earlier years poor SLK data was more common, with records from Queensland having a high rate of missing name data in 2005–06—the first year it reported unit record data to the ACAP MDS. After using reported suburb to derive postcode for cases where it was missing, postcode data used in data linkage was missing for less than 1% of clients in all data sets.

In general, SLK data was less likely to be missing or unreliable on the ACAP MDS than on the HACC MDS. Records with missing elements of SLK-581 were less common on the ACAP MDS than on the HACC MDS from 2009–10 onwards, although in all years fewer than 0.5% of HACC quarterly records had insufficient data for linkage. However, assuming that people are equally likely to be born on any day of the year, 1 January birthdays were 8 to 15 times more common than expected on the HACC MDS (uniform birthdays implies a proportion of $1/365 = 0.27\%$ birthdays on any one day of the year). This compares with 2 to 3 times more likely on the ACAP MDS. As mentioned in Section 2, the high prevalence of 1 January birthdays on the HACC MDS resulted in the special treatment of these C_CIDs; such C_CIDs were included when matching HACC data sets to records with full name information but were not included in the PIAC database otherwise.

For PIAC 2014, the data are of similar quality.

3.4 Linkage quality

Name-based linkage

When undertaking name-based linkage, two measures of quality were estimated: the positive predictive value (PPV), which is the proportion of match pairs that are true matches, and sensitivity which estimates the proportion of true matches identified. These 2 measures are combined, using a geometric average, into an F-score to allow simple comparisons across different name-based strategies.

Using the above approach, the name-based linkages undertaken when deriving the PIAC database were estimated to have F-scores of at least 99%.

Key-based linkage

There are many combinations of the available key components that could be used to define match keys. To ensure that any employed match keys were based on combinations of components that both discriminated well between individuals and would not introduce too

many false positives, 3 measures—calculated for each match key—are used to identify suitable linkage keys and their order of use:

- the *estimated false match rate* (FMR) for links established using the match key (the lower the better)
- the *estimated marginal trade-off* (m_tf) between additional true and additional false matches for links established using the match key when compared with matches made by a slightly more precise key (the higher the better)
- a measure of *discriminating power* (expressed as a percentage). This is the product of the unique key rates for the 2 data sets being linked, where the unique key rate is the proportion of records within a data set that have a unique value for the key in question (the higher the better).

The first 2 of these are used to identify keys to be used in the linkage process by setting cut-offs, while the third determines their order of use (highest to lowest). For PIAC 2014, the KBL processes use an FMR limit of 0.5% and an m_tf lower limit of 5 to select suitable keys.

To assess the quality of the linkage, using PIAC 2011 data a number of comparisons were undertaken using name-based and key-based linkage to link data sets with full name information. In the comparisons, the KBL processes used only SLK-581 and postcode. Using the name-based linkage as the reference, the PPV ranged between 95.6% and 98.5% for KBL, and sensitivity was between 90% and 91%, giving F-scores of 93% to 94%. Processes using additional data—for example, when matching SPARC to ACAP the date of assessment was also used—are expected to have both higher PPV and sensitivity (AIHW 2014).

Additional match process for ACAP MDS

Noting that the NACDC table ACCR_ASSESSMENT has information on complete assessments brought across from the ACAP assessment form (ACCR form) and that this is the same source as the ACAP MDS data, before linking ACAP to the master list using KBL as described above, ACAP MDS data for complete assessments with at least 1 approval were linked deterministically to SPARC and CACP clients on the ACCR_ASSESSMENT table using a small number of selected keys. This was done to improve the quality of matching by associating ACAP MDS records with named data before matching more generally to the master list. This matching used keys consisting of full SLK-581, recipient postcode data and data items common to the ACAP MDS and the ACCR_ASSESSMENT tables, including: ACAT ID; approval indicator variables; assessment date variables; and some demographic variables. Seven keys were used in the matching process; all included at least SLK-581 and full postcode, and over 99% of matches were made using keys that also included ACAT ID in the key.

4 Program and data issues affecting analysis

There have been a number of changes to policy and data collection instruments over the years which have affected the data being collected and stored on NACDC. Some of these changes affect how data should be derived for particular analyses, while others affect the interpretation of results. The main changes are listed below.

Examples of how differences in reporting practises and varying data quality can affect the derivation of analysis data sets can be seen in Section 5.

The NACDC data supply process is described in more detail in Section 3 of the NACDC Data Dictionary.

4.1 Residential aged care

Reforms due to the Aged Care Act 1997

Nursing homes and hostels were amalgamated into a single system during 1997–98 under the *Aged Care Act 1997* (AIHW 1999). Data in NACDC on the use of residential care are from this period onward.

With the unification of the 2 sectors into ‘residential aged care’, the eight-category Resident Classification Scale (RCS) was introduced on 1 October 1997. Income-tested RAC fees started on 1 March 1998.

Later aged care reforms

From 1 July 2014, the distinction between low care and high care in PRAC was removed as part of the 2012 aged care reforms.

A new fee structure applied to residents who enter permanent residential care from 1 July 2014. This change led to a short-term increase in people entering PRAC just prior to 1 July 2014, and a corresponding fall afterwards, to avoid the inclusion of assets—and not just income from assets—in the means test: all new residents could still be asked to pay a basic daily fee of up to 85 per cent of the single basic age pension, while some residents could be asked to pay an additional means-tested care fee based on an assessment of both their assets and income.

Funding tools for permanent residential aged care

Resident Classification Scale (RCS)

Within PRAC, since 1997 care needs of residents have been recorded by a number of appraisal instruments in order to determine subsidy levels. Pre-1997, the funding tool was called Resident Classification Instrument and used in nursing homes. These data are not available in NACDC.

The Resident Classification Scale (RCS) V1.0 began operating on 1 October 1997 and covered both nursing homes and hostels to assess people for funding purposes. It determined the level of subsidy to which facilities were entitled for each resident (based on people’s levels of dependency). RCS appraisals were conducted annually for each resident.

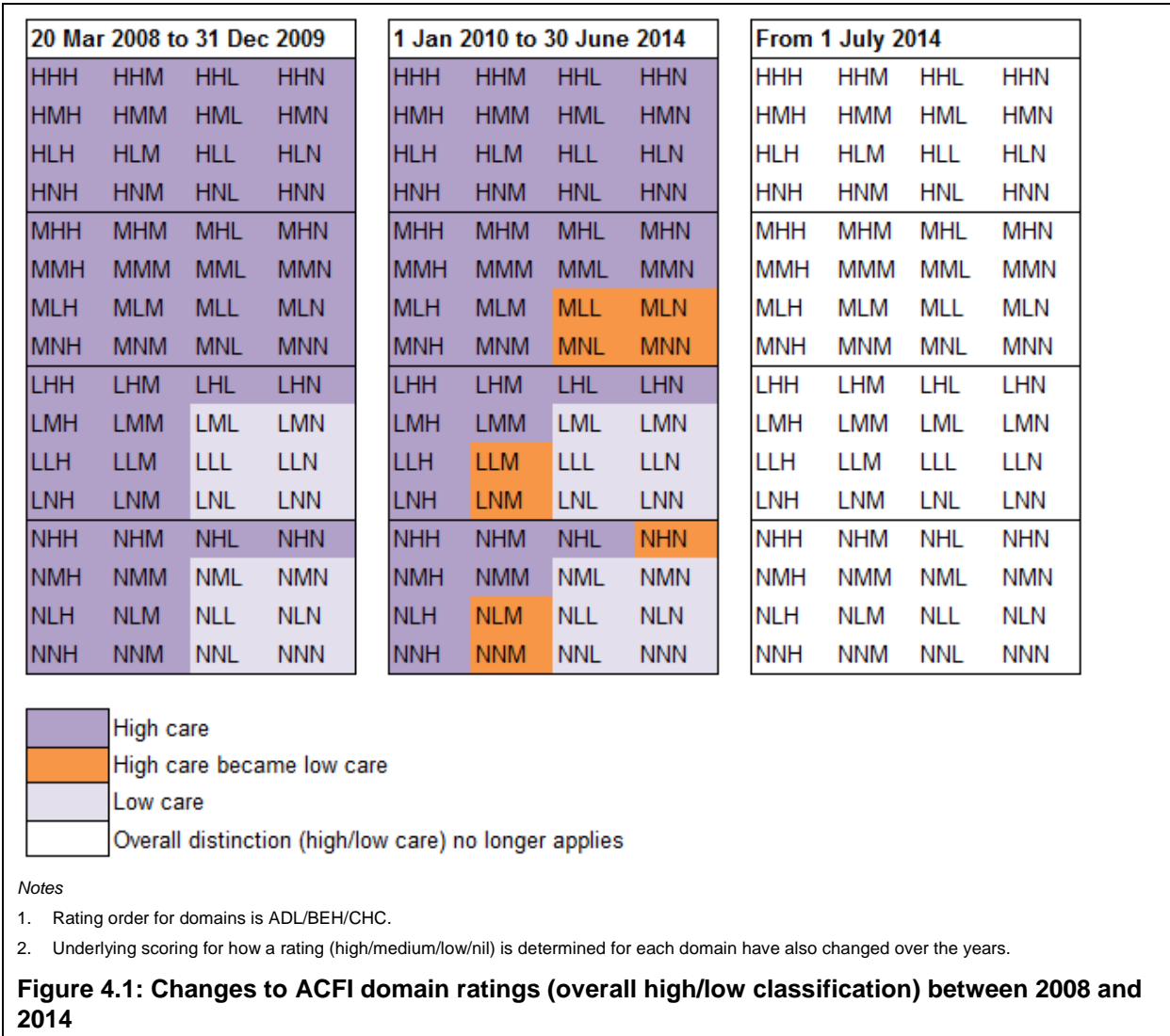
Rejected appraisals (for example, those not accepted for funding purposes) are not included in the NACDC tables.

The RCS V1.0 was revised during 1998, and RCS V2.0 was brought in on 1 November 1998. The number of questions was reduced from 23 to 21, and several questions were adjusted. The RCS was grandfathered from March 2008, with all new assessments conducted under the Aged Care Funding Instrument.

Aged Care Funding Instrument (ACFI)

ACFI was implemented on 30 March 2008 (ACFI) to replace the RCS. Appraisals are conducted on an ‘as-needed’ basis (that is, when care needs have changed to the extent that the subsidy is affected), rather than annually. ACFI assesses people’s care needs against 3 care domains—activities of daily living (ADL), behavioural (BEH), and complex health care (CHC). Levels of subsidy are based on different rankings (low, medium or high) in these 3 domains, and again, rejected appraisals are not included in NACDC data.

In January 2010, the definition of high- and low-care in terms of the 3 care domains measured by the ACFI was adjusted, and in July 2014, the overall high/low classification was removed entirely (Figure 4.1).



The removal of the distinction between low care and high care in PRAC (see issues under ACAP) may have reduced the need for re-appraisal for people with care needs bordering between low- and high-care.

The interpretation and scoring of questions over time is best gauged by looking at the relevant RCS and ACFI user guides. Descriptions of reporting codes for particular questions are given in the detailed documentation of NACDC tables.

4.2 Transition Care Programme

Data items on the use of TCP are recorded in the same NACDC table as admissions to RAC, EACH and EACHD. Additional data items are available for TCP, including information on the functional status of TCP clients at the start and end of the care episode. Comparable measures are not available for the other programs included in PIAC.

4.3 Home care packages

The largest of the pre-2013 programs, Community Aged Care Packages (CACP), was operational from 1992, Extended Aged Care at Home (EACH) from 2002, and Extended Aged Care at Home—Dementia (EACHD) from 2006. NACDC holds reliable data for CACP from 1997–98 onwards, and from the beginning of the operation of the other 2 programs.

In 2013, CACP, EACH and EACHD were replaced by the Home Care Packages Program (HCP). From a data perspective, the main impact is that HCP use is recorded in an additional set of tables in the NACDC that are not readily integrated with the SPARC/CACP tables used previously for these programs. The tables in NACDC used to record HCP use are structured somewhat differently from those used to record the use of the earlier package programs.

However, both sets of tables are broadly similar in that program data capture episodes of care, and each person in care is also identified by a unique recipient identifier. However, while the recipient identifiers used for the earlier packages have carried over into HCP where people were in care at the time of the transition, people who had left CACP/EACH/EACHD and later returned to HCP are not reliably identified in the data, as these people may have later received a new recipient identifier for HCP.

In addition, a person and their package use may appear in both sets of tables. This particularly affects people who migrated to the new program from CACP, EACH or EACHD. This means that care must be taken to ensure that program use events are not counted twice when bringing them together using the PIAC link map. Information that is particularly affected by possible duplication includes:

- number of clients at a point in time or over a period
- admissions events (that is, entry and exit dates into a program)
- leave events (when a person goes on leave from program use; for example due to hospitalisation or visiting family).

Both sets of tables were used to create the PIAC 2014 link map, but in the future this may change. The HCP tables include historical data on CACP/EACH/EACHD, and are generally considered more comprehensive and reliable, as these historical data are updated in each annual supply of NACDC data—the CACP/EACH/EACHD tables are no longer being updated.

On 1 July 2014, new income testing arrangements became effective for care recipients entering into a home care agreement from that date, with income tested care fees being

applied consistently to ensure that people with similar income pay similar fees. People entering HCP from 1 July 2014 can still be asked to pay a basic daily fee of up to 17.5 per cent of the single basic age pension. In addition, if their income is over a certain amount, they can expect to be asked to pay an income tested care fee, which reduces the amount of care subsidy paid by the Government. Safeguards include annual caps and lifetime caps on the income-tested care fees payable by care recipients. These new fee arrangements could affect the number of people taking up packages.

4.4 Program leave

People in PRAC or on a home care package (EACH, EACHD, CACP, HCP) can go on 'leave' for a number of reasons. That is, they temporarily leave their accommodation in a facility or their usual residence in the community where they are receiving community services through a package. For these people, their place in the care program is retained until they return from their leave and the government subsidies are not affected (within limits). People in RRAC cannot access leave.

Data on leave episodes are provided on two tables in NACDC: one for PRAC, EACH, EACHD and CACP, and one for HCP.

Leave is reported using different code sets for different programs. More specifically, different sets of codes are used for people on CACPs compared with people on EACH, EACHD or in PRAC, while the leave table for HCP includes leave descriptions rather than leave codes. Many periods of leave from a package are reported on both NACDC source tables. Duplicate records should be dropped. The record with the better description of reason for leave should be retained:

- HCP description is preferred over the CACP description.
- RAC/EACH/EACHD description is preferred over the HCP version.

By comparing the duplicates, 'Alternative Care Leave' for CACP clients seems to equate to HCP leave for 'Respite Care', and 'Other Care' for CACP clients seems to relate to HCP 'Social Leave'.

Not all leave is reported. A previous linkage project found that about 20% of hospital leave events from PRAC were not reported to the subsidy payment system. This potentially reflects the lack of incentive to report—short leave periods, for example any leave below 30 days for PRAC, do not affect subsidies. See Appendix B of *Movement between hospital and residential aged care 2008–09* (AIHW 2013) for more information.

4.5 Aged Care Assessment Program

Before a person can access RAC, TCP or a home care package, they must obtain an approval following an assessment under the ACAP conducted by an Aged Care Assessment Team (ACAT; also known as Aged Care Assessment Service or ACAS). Note that an ACAT approval for a particular type of care cannot be given without the consent of the client.

Key changes to ACAP assessment include:

- from 1 July 2004, residents moving between low- and high-care within the same aged care home no longer needed an ACAP assessment
- up to 30 June 2009, an ACAT approval remained valid for 12 months; that is the ACAP client had up to 12 months to take up approved care

- on 1 July 2009, the time limit on take up approval was removed for many care types: residential respite care (high- and low-care), high-level PRAC, EACH and EACHD packages (that is, high-level home care packages). Consequently, all approvals made for these care types on or after 1 July 2008, which were not explicitly time limited, did not lapse even if the person had not received that care
- up to 30 June 2014, an ACAT approval was required for residents moving between facilities in order to change from low care to high care. From 1 July 2014, this requirements was removed, along with the distinction between low care and high care in PRAC as part of the 2012 aged care reforms.

Unit record data for all assessments, both completed and incomplete, are available. An assessment (as identified using the `assessment_ID`) can be recorded multiple times on the NACDC table, primarily due to updated information. For analysis, just one record per assessment should be selected (see section 5.3). In general, the last update of the assessment information is to be preferred. For other information on a method for determining the preferred record for an assessment see Section 5.3.

There are some inconsistencies in the data (for example, dates, duplicates, overlapping assessments by the same ACAT for the same person). People may also have concurrent assessments by different ACATs. An approach for date cleaning is described in Section 5.3.

Variables on program use at the time of an assessment is provided in the ACAP MDS. The AIHW experience has been that people may not know through which program their assistance is provided. Data on government programs being accessed at the time of assessment can also be derived from the special analysis files

`SERVICE_EVENTS_AND_DATES` or `CARE_PATHWAYS` tables described in Section 5 (these do not include VHC, nor Day Therapy, National Respite for Carers or Carelink services; along with HACC these 3 became part of the Commonwealth Home Support Program).

4.6 Home and Community Care

Before 2012, HACC was provided by states and territories, and funded by the Commonwealth (Australian) government, and the program catered to a broad population, such as younger people with disability as well as frail older people. On 1 July 2012, the Australian Government assumed full policy, funding, and day-to-day responsibility for HACC services for people aged 65 and over, and for Aboriginal and Torres Strait Islander people aged 50 and over in all states and territories except Victoria and Western Australia (this new program was termed 'Commonwealth HACC' to distinguish it from the 'old' HACC services which continued to operate in Victoria and WA). Services for those aged under 65 were to be funded under the state responsibilities for disability. This could have impacted the quality and coverage of the HACC MDS.

The Commonwealth HACC then became the main part of the Commonwealth Home Support Program (CHSP) from July 2015. This is also likely to impact the coverage of the 'HACC' MDS. The data collection method also changed, and the impact these had on the data, or its quality, was not known at the time of this guide's publication. Complete CHSP data is not expected to be available for the first year of the program's operation.

HACC service use data are reported to the HACC MDS showing the services provided to a client in a quarter by an agency. In general, the number of records a person has in a quarter indicates the number of providers they are getting services from:

- In a small percentage of cases agencies supplied multiple records for a person in a quarter. It is not clear why this was the case, but differences in reported dates, volume of service and demographics have been seen.
- The above 'duplicate' records only affect the derivation of periods of HACC use—as derived for the cohort pathways file—if the duplicates were the result of corrections to service provision because some services were erroneously reported as provided.
- In a small percentage of cases agencies supplied multiple records for a person in a quarter. It is not clear why this was the case, but differences in reported dates, volume of service and demographics have been seen.
- The above 'duplicate' records only affect the derivation of periods of HACC use—as derived for the cohort pathways file—if the duplicates were the result of corrections to service provision because some services were erroneously reported as provided.

Data on services provided by HACC were collected via the HACC MDS V1.0 for 2001–02 to 2004–05 and HACC MDS V2.0 from 2005–06. There are some differences between the two versions:

- There were some changes in demographic variables.
- Functional status variables were added for HACC MDS V2.0. However, they are not well reported, with around two-thirds of records having missing values.
- Service use start and end dates were introduced in HACC MDS V2.0. However, these are provider-specific and are not well-reported. Also, the dates do not relate to specific service types.
- There were some changes in the categories of service use used for reporting.
- The variables REASON_FOR_HACC_CLIENT_STATUS and ACCOMMODATION_AFTER_CESSATION relate to HACC MDS V1.0 only (for example, pre-2005).
- MDS V1.0 collected data on care recipients. One service use type—respite care—was considered to be for the carer and so was meant to be recorded under the carer's SLK-581 (that is, the carer was the recipient of the service). This does not seem to have been done consistently.
- In MDS V2.0, service use data were reported for care recipient and carer pairs, with SLK-581 being reported for both people and services for either person being reported in a 'pair' record.

Other specific problems with data quality have also been identified:

- Carer dates of birth are not very reliable: 1 January seems to have been used as a default and the start of the decade birthdays are also overly common.
- Reported service volumes are sometimes unreasonably large. Volumes that are too large can be identified using the following (rather generous) cut-offs:
 - valid hourly services are ≤ 2000 per agency per quarter
 - valid item services are ≤ 300 per agency per quarterand
 - cut-off applies to \$ value of home modifications.

Only HACC V2.0 are on the NACDC. Data for HACC V1.0 are held separately in a number of tables (sometimes one file per year, sometimes 4 files per year).

5 PIAC special analysis data sets

Person-based analysis data sets can be derived by bringing together the data relating to individual p_ids. When bringing the data together, inconsistencies may be observed in the integrated data. There are two main reasons for this:

- the source data sets have a range of purposes and
- data about a client may have been provided by different people over an extended period.

The approaches taken to resolve inconsistencies for two types of variables are described below: stable demographic variables, which should not change over time, and program use dates. In addition, several ways of turning a set of program use dates into a care pathway are presented.

5.1 Demographic characteristics

There are a number of client demographic characteristics that should not change over time; for example date of birth (DOB), date of death (DOD), sex and country of birth. However, sometimes different data sets may contain different values. In addition, there may also be variation within data sets where demographic values are reported by a number of different service providers or on separate occasions. The ACAP and HACC MDSs fall into this latter category.

The PIAC 2011 rules used for getting preferred demographics have again been used, except for DOD, see Appendix B6 of *Patterns in Use of Aged Care 2002–03 to 2010–11* (AIHW 2014). In essence, a majority rule is used within and across programs reporting a demographic variable. If there is a tie within a program the date of last use is used to break the tie. Across programs, if there is a tie, then a program priority ranking is used to break the tie. For the latter, HCP's priority rank is between that of CACP and ACAP.

All DOD data come from the NDI. However, due to variation in reporting on the NDI, in a small number of cases more than one DOD is associated with a p_id. The preferred DOD is determined by comparing DODs for a p_id with last service program use, and is based on:

- DOD closest to last service end date (absolute difference)
- for ties:
 - whether the record has a valid mortality database mort_id that can be used to bring in cause of death (if approved)
 - the most recent record (for example, highest NDI record_id) if either the ties both have a mort_id or neither have a mort_id.

Note that for 5 p_id linked to an NDI record the DOD is missing. Note, also, there are 62 p_id on the events file SERVICE_EVENTS_AND_DATES (see Section 5.2) that are not on the demographics file. These people had 71 HCP service use events (and no other events), of which just 3 were before 1 July 2014. The reason for the 'loss' of these p_id on the demographics file is that there is no demographic information for the HCP client ID on the relevant HCP NACDC15 table **AND** the p_id did not link to any other program or death. The data for linkage for these people came from the NACDC for SPARC and CACP clients via the HCP–SPARC–CACP mapping file on NACDC.

Reference table PIAC_PREFERRED_DEMOGS_DOD

The table with preferred demographics to be used in analyses contains one record per p_id on the PIAC linkage map (5,028,933 records). All non-date variables use descriptive codes that can be easily turned into one-digit classification codes (if preferred). There are 16 variables in the table (Table 5.1).

Table 5.1: 'Master' demographic variables for PIAC 2014

Name	Type	Length	Format	Description
p_id	Char	10	\$10.	PIAC person identifier
dob_mstr	Num	8	DATE9.	Date of birth
DOD_mstr	Num	8	DATE9.	Date of death
sex_mstr	Char	14	\$14.	Sex
ind2_mstr	Char	25	\$25.	Indigenous status, revised via country of birth and language
Abl_mstr	Char	17	\$17.	Aboriginal status
TSI_mstr	Char	18	\$18.	TS Islander status
ind_ever	Char	20	\$20.	Person ever recorded as Indigenous, via ind2_mstr/program
ind_always	Char	20	\$20.	Always indigenous, via ind2_mstr/program
ind_major	Char	20	\$20.	Majority indigenous, via ind2_mstr/program
ever_ind_ever	Char	20	\$20.	Person ever recorded as ever_indigenous, via any program
cob_mstr	Char	21	\$21.	Country of birth
ep_mstr	Char	10	\$10.	English proficiency
lang_mstr	Char	24	\$24.	Preferred language
num_progs	Num	8		Number of programs linked to
mort_id_mstr	Char	15	\$15.	Mortality identifier for NDI

Note: English proficiency (EP) is derived from country of birth. For a full explanation of the process, see Box 3.1 in *Patterns in Use of Aged Care 2002–03 to 2010–11* (AIHW 2014).

5.2 Service use dates

Event dates are reported for different purposes in different ways and by different people on the various data sets. Factors affecting reporting of program use dates include:

- in the program data sets with full name data, dates are reported as part of payment systems
- the dates on the ACAP and HACC MDSs are used to report on general service provision and program use, and are not used directly for funding purposes
- HACC program use is reported quarter by quarter by service providers
- service use dates for people using home care packages may sometimes be recorded in more than one table in the NACDC (depending on whether the package in question is HCP, or its predecessors CACP, EACH or EACHD), leading to duplication of service use records
- in some cases—especially when services are provided in the community—exact start and end dates of service use may not be known by those reporting the dates.

Because of the above, program use dates that imply concurrent use of two incompatible programs may be reported, and reported dates may not be consistent with a person's date of death. For these reasons, a range of edits are applied to reported service use dates. The same edit rules as those used for PIAC 2011 are again used, with DOD being used to truncate all program use dates for a p_id. For HCP, the rules applied to CACP dates are used for HCP levels 1 and 2, and for HCP levels 3 and 4 the rules applied to EACH and EACHD dates are used. These rules are described in Appendix B6 of *Patterns in Use of Aged Care 2002–03 to 2010–11* (AIHW 2014). The final event dates for all programs can be used in both analyses of program use over time, and use of programs at a point in time.

The edit rules can result in truncation, splitting or deletion of service use events. After applying the edits, within each program (except HACC), the number of events was within 1% of the original number. For HACC, the number of events was slightly larger due to the combined effects of a less accurate process for reporting service use dates, and because many HACC services should not be used at the same time as PRAC or home care packages.

When estimating the number of clients a program has during a period, a person should be assumed to be receiving services on the date they start with a program but not on the day they leave a program. This avoids double counting of clients. Details of the processes and rules applied to derive the final set of event dates are described Appendix B6 of *Patterns in Use of Aged Care 2002–03 to 2010–11* (AIHW 2014).

The events table contains periods of service use by the care recipient. Consequently, HACC service events on this table *exclude* time spent by HACC service providers on case management and planning. Also excluded are services provided to carers (respite/carer counselling). This means that there are p_ids on the HACC concordance and demographics tables that are not on the event table.

Reference table SERVICE_EVENTS_AND_DATES

The events table contains one record per service use event or death (14,074,623 records) (Table 5.2). It does not contain ACAP assessments. These are given in a separate table (see below). People that only had ACAP assessments are therefore not included in this table.

Table 5.2: Service use variables for PIAC 2014

Name	Type	Length	Format	Description
p_id	Char	10	\$10.	PIAC person identifier.
SERVICE_ID	Char	20	\$20.	This is the service provider ID (as on NACDC15) for RAC, CACP and HCP events. It allows the identification of transfers.
end_date_revised	Num	8	DATE9.	Service use end date after revision for incompatible overlap. For deaths, this is the same as start_date_revised.
event	Char	2	\$2.	Service (program) type used, with HACC recorded as 'N' (nursing and allied health), 'B' (Centre-based day care) or 'O' (other), as used, to identify inconsistent overlap.
event2	Char	2	\$2.	Service (program) type used, with all HACC events recorded as 'H'.
start_date_revised	Num	8	DATE9.	Service use start date after revision for incompatible overlap. For deaths, this is the same as the end_date_revised.
status	Char	4	\$4.	Whether p_id has a death event (recorded as 'dead' or blank). For the 5 p_id with poor DOD information, the event is included in the table, but the dates are missing.

A single character is used to specify the service (program) type used in the variables event and event2 (Table 5.3).

Table 5.3: Characters used in variables event/event2 for PIAC 2014

Character	Description
*	Death
P	PRAC
R	RRAC
C	CACP
E	EACH
D	EACHD
T	TCP
2	HCP levels 1–2
4	HCP levels 3–4
H	HACC (N, B or O)
N	HACC Nursing/allied health
B	HACC centre-based day care
O	HACC Other (not N or B)

5.3 ACAP assessments with cleaned dates

A person may get an assessment under ACAP at any time. Consequently, there is no need to adjust assessment dates because the person is receiving another service during the assessment period. However, within p_id there are some inconsistencies in reported ACAP assessment dates, and a number of edits are undertaken to resolve these. In addition, inspection of the ACAP data sets shows that sometimes assessments are included more than once, and so all duplicates need to be removed. For this table, the most recent version (based on the NACDC ACAP MDS variable effective_start_date) is chosen as the preferred record.

It is possible for a person to be getting two assessments at the same time. However, a person should not have two assessments happening at the same time with the same ACAT. There are cases in the ACAP MDS data of people apparently receiving two assessments by the same ACAT at the same time. A number of rules were therefore applied to resolve the start and end dates for these overlapping assessments. The rules are based on the nature of the overlap and whether the two assessments in question were complete (that is, had sign-off on decisions been made on approvals) or were closed off while incomplete. Assessments are dropped if, after applying these rules, the start date is after the end date. The rules used to adjust are described in Appendix B6 of *Patterns in Use of Aged Care 2002–03 to 2010–11* (AIHW 2014). Only retained assessment records were used when deriving the preferred values for demographic variables within ACAP.

In addition to the above, the preferred DOD (see above) is used to truncate the assessment end date and to drop any assessments starting after the preferred death date. Truncation of assessments at DOD assists with the derivation of the care pathways (see Section 5.4). However, assessments being undertaken at the time of death may have been closed by the ACAT at a later date due to delays in communication of the fact of death.

Reference table ACAP_MDS_CLEANED_DATES

The table contains one record per ACAP assessment on the MDS, after removing duplicate records and revising start and end dates for inconsistency within and across assessments (2,258,761 records). Note that the ACAP MDS contains other assessment dates (for example, first face-to-face-contact) and these other dates reported for each assessment have not been adjusted here.

This table contains the preferred record for an ACAP assessment. The PIAC p_id is also provided so that it can be easily integrated with other tables (Table 5.4).

Table 5.4: ACAP MDS variables for PIAC 2014

Name	Type	Length	Format	Description
p_id	Char	10	\$10.	PIAC person identifier
ASSESSMENT_ID	Char	40	\$40.	ACAP assessment identifier (ASSESSMENT_ID as on NACDC15 table)
EFFECTIVE_START_DATE	Num	8	DATETIME20.	The date the record is effective from (EFFECTIVE_START_DATE as on NACDC15 table ASSESSMENT_DETAILS_ACAP)
ACATID	Char	3	\$3.	Identifies the ACAT team which conducted the ACAP assessment (ACATID as on NACDC15 table ASSESSMENT_DETAILS_ACAP)
record_id	Char	52		Record identifier (record_id = compress (assessment_ID effective_start_date))
completion_status	Char	1		Assessment completion status (A = complete assessment, I = incomplete assessment)
ACAP_end_date	Num	8	DATE9.	ACAP event end date, after cleaning
ACAP_start_date	Num	8	DATE9.	ACAP event start date, after cleaning

Note that because there is only one record per assessment ID, the derived dates can be added to the detailed assessment information using ASSESSMENT_ID only. However, EFFECTIVE_START_DATE indicates the preferred record.

5.4 Care pathways

As people's care needs change, the services they access can also change. The various changes in program use is called here a 'care pathway'. There are many ways to look at care pathways. Pathways using several different representations have been derived. All of these present the pathways as a list of events in chronological order, with each type of event represented by a particular character. The longest care pathway for PIAC 2014 has 110 care program events and/or ACAP assessments.

The care pathways were derived using the service event and cleaned ACAP MDS assessment files (see sections 5.2 and 5.3). The event order is based on the start date of the event. Note that because (a) people can have ACAP assessments while accessing a care program, and (b) people can be clients of more than one program at the same time, the last event in the pathway might not be the event that ended last (for example, with the most recent end date). Also, there are 4 care pathways with missing event start and end dates. These pathways consist solely of a death which had missing date of death data.

Including both program use and ACAP events, for PIAC 2014 there are 4,796,550 people with a care pathway. This is less than the number of people in the demographics file for two reasons:

- HACC service events *exclude* time spent by service providers on case management and planning. Also excluded are HACC services provided to carers (respite/carer counselling). Consequently, people that only had these HACC services (for example, no other program use and had not died) are excluded from care pathways.
- There are 20,407 p_id who were HCP clients only who do not appear on the care pathways file. These HCP clients were all ACAP assessment-only clients (for example, no admission), and most of these assessments related either to 2014–15 (for example, were out of scope for PIAC 2014) or were before the full implementation of the ACAP MDS. In all, there were 388 HCP clients with an in-scope assessment that had not matched to an assessment on the ACAP MDS (for example, a missed match) and so these are not represented on the care pathways file.

Reference table CARE_PATHWAYS

The table contains one record per person (p_id) (4,796,550 records). Six different care pathway variables are included, each presenting the care pathway slightly differently (Table 5.5): straight_path; simple_path; path_repeats; path_skel; path_care; and path_care_changes.

The first four pathway types include both service and ACAP events, while ACAP events are excluded for the last two.

Also included in the reference table are a number of descriptive variables relating to the order of HACC and ACAP use, as well as the start and end dates for each event in the pathway (Table 6).

Table 5.5 (continued): Pathways variables for PIAC 2014

Name	Type	Length	Format	Description
p_id	Char	10	\$10.	PIAC person identifier
first_A	Num	8		Event number for first completed ACAT
first_H	Num	8		Event number for first HACC use
first_PRCDET	Num	8		Event number for first ACAT-program use
first_hassessment	Num	8	DATE9.	Date of first HACC assessment
HACC_before_ACAP	Char	35		Describes relationship between 1st HACC and ACAP
possible_missing_ACAT	Char	3		Indicates if ACAP assessment was possibly missing from MDS; e.g. before the 2005/2 quarter.
straight_path	Char	110		Full path, simple concatenation of events (with ACAT)
simple_path	Char	110		Full path, simple concatenation of events with transfers indicated (with ACAT)
path_repeats	Char	61		Path, like straight_path but indicating repeated event types but not the number of repeats(with ACAT)
path_skel	Char	44		Path showing only care changes (with ACAT)
path_care	Char	30		Path showing only care order (no ACAT)
path_care_changes	Char	104		Path showing non-transfer care changes (no ACAT)
Ps	Num	8		Total P-P transfers (same or next day)
Rs	Num	8		Total R-R transfers (same or next day)

(continued)

Table 5.5 (continued): Pathways variables for PIAC 2014

Name	Type	Length	Format	Description
date_in1	Num	8	DATE9.	First event start date
date_out1	Num	8	DATE9.	First event end date
date_in2	Num	8	DATE9.	Second event start date
date_out2	Num	8	DATE9.	Second event end date
date_in3	Num	8	DATE9.	Third event start date
date_out3	Num	8	DATE9.	Third event end date
date_in108	Num	8	DATE9.	108 th event start date
date_out108	Num	8	DATE9.	108 th event end date
date_in109	Num	8	DATE9.	109 th event start date
date_out109	Num	8	DATE9.	109 th event end date
date_in110	Num	8	DATE9.	110 th event start date
date_out110	Num	8	DATE9.	110 th event end date
status	Char	5		Indicates whether the person has a death record on PIAC 2014 ('dead' vs 'alive')

As in the service events file in Section 5.2, in the care pathways a single character is used to specify an event (tables 5.6 and 5.7).

Table 5.6: Characters used in variables straight_path/simple_path/path_repeats/path_skel/path_care/path_care_changes for PIAC 2014

Character	Description
*	Death
P	PRAC
R	RRAC
C	CACP
E	EACH
D	EACHD
T	TCP
2	HCP level 1–2
4	HCP level 3–4
H	HACC
A ^(a)	Complete ACAP assessment, that <i>started</i> when the person was NOT in RAC
a	Complete ACAP assessment, that <i>started</i> when the person was in RAC
I	Incomplete ACAP assessment, that <i>started</i> when the person was NOT in RAC
i	Incomplete ACAP assessment, that <i>started</i> when the person was in RAC
T ^(b)	Same or next day transfer for care (not assessment or HACC) of the same type (e.g. 'Pt' indicates that a person in PRAC transferred to another facility to continue their care, 'Pat' indicates that a person in PRAC transferred to another facility to continue their care after having an ACAP assessment while in the first facility)
X	Indicates that a type of event (care or assessment) repeated more than once with no other care type in between (for example, 'Rx' indicates that the person had multiple non-transfer RRAC events in a row; 'Patx' indicates that a person in PRAC transferred to several other facilities in succession after having an ACAP assessment while in the first facility)

(a) Note that the letters 'A', 'a', 'I', 'i' had different meanings (still related to assessments) in the care pathways derived for the PIAC 2006 analysis.

(b) Because people can get different services through different HACC agencies at the same time, concurrent and adjacent HACC events were combined before deriving care pathways.

Table 5.7: Examples showing the relationship between the different types of variables available for 'pathways'

straight_path	simple_path	path_repeats	path_skel	path_care_changes	path_care
HRRRRPP*	HRRRRPt*	HxRxPt*	HRP*	HRRRRP*	HRP*
HAIP*	HAIP*	HAIP*	HAIP*	HP*	HP*
AP*	AP*	AP*	AP*	P*	P*
AHDRARP*	AHDRARP*	AHDRARP*	AHDRARP*	HDRRP*	HDRP*
HAAP*	HAAP*	HxAP*	HAP*	HP*	HP*
RARRP*	RARRP*	RARxP*	RARP*	RRRP*	RP*
HPa*	HPa*	HPa*	HPa*	HP*	HP*
HAIPP*	HAIPt*	HAIPt*	HAIP*	HP*	HP*
HHHHACAHRP*	HHHHACAHRP*	HxACAHRP*	HACAHRP*	HHHHCHRP*	HCHRP*
PPP*	PPt*	Pxt*	P*	PP*	P*
HPP*	HPt*	HPt*	HP*	HP*	HP*
HAHHHAPP*	HAHHHAPt*	HAHxAPt*	HAHAP*	HHHHHP*	HP*
HHHHHAAIRRARPI*	HHHHHAAIRRARPI*	HxAxIRxARPI*	HAIRARPI*	HHHHHRRRP*	HRP*
AHRCARaERRP*	AHRCARaERtP*	AHRCARaERtP*	AHRCARaERP*	HRCRERP*	HRCRERP*
HIAPP*	HIAPt*	HIAPt*	HIAP*	HP*	HP*
PPPP*	PPtt*	Pxtx*	P*	PP*	P*
AIAPP*	AIAPt*	AIAPt*	AIAP*	P*	P*
PPPP*	Pttt*	Ptx*	P*	P*	P*
HAAAAPPAPP*	HAAAAPtAPt*	HxAPtAPt*	HAPAP*	HPP*	HP*
PPPP*	Pttt*	Ptx*	P*	P*	P*
APIATE4*	APIATE4*	APIATE4*	APIATE4*	PTE4*	PTE4*
AHCRPa*	AHCRPa*	AHCRPa*	AHCRPa*	HCRP*	HCRP*
HHHHHHACACCHH A4*	HHHHHHACACtHHA 4*	HxACACtHxA4*	HACACHA4*	HHHHHHCCCHH4*	HCH4*
ACACH2	ACACH2	ACACH2	ACACH2	CCH2	CH2
RPPPP*	RPttt*	RPtx*	RP*	RP*	RP*
HAAHHHARPRARaP PPRR4	HAAHHHARPRARaPt tRR4	HxAxARPRARaPt Rx4	HAHARPRARaPR4	HHHHRRRRPRR4	HRPRPR4
HHHIAIIIRHIAPaaP	HHHIAIIIRHIAPaat	HxIAIxARHIAPaxt	HIAIARHIAPa	HHHRHP	HRHP
HHAAPaPPP*	HHAAPattt*	HxAxPatx*	HAPa*	HHP*	HP*
AHAAPPa	AHAAPta	AHxApta	AHAPa	HP	HP
AAARAAHA4	AAARAAHA4	AxRxAHA4	ARAHA4	RH4	RH4

7 Accessing PIAC data and information

Access to PIAC 2014 data and information can be facilitated in a number of ways. AIHW releases a range of reports and publications that use the linked data; these are available for free download from the AIHW Data linkage website. Publications relating to aged care are also freely available through GEN, the AIHW aged care website.

PIAC is a valuable resource and access can be facilitated through both customised data requests and through the integration of additional datasets to create an expanded linked database. The AIHW is an accredited Integrating Authority, and is able to facilitate both these avenues for researchers.

All data linkage projects and access to AIHW linked data can only be undertaken with the approval of the independent AIHW Ethics Committee.

Further information on the AIHW data processes is also available through the AIHW Data governance framework.

Appendix A: Data Quality Statements

Relevant Data Quality Statements are available:

- DQS for the [National Aged Care Data Clearinghouse \(NACDC\)](#);
- DQS for the [Aged Care Funding Instrument \(ACFI\)](#); and
- DQS for the [National Death Index \(NDI\)](#).

References

- AIHW (Australian Institute of Health and Welfare) 1999. Australia's welfare 1999: Services and assistance. Australia's welfare series no. 4. Cat. no. AUS 16. Canberra: AIHW.
- AIHW 2006. Data linkage and protecting privacy: a protocol for linking between two or more data sets held within the Australian Institute of Health and Welfare. Canberra: AIHW. Viewed 30 September 2014, <<http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?ID=6442468416>>.
- AIHW 2009. Pathways through aged care services: a first look. AIHW bulletin no. 73. Cat. no. AUS 116. Canberra: AIHW.
- AIHW 2010. Dementia and the take-up of residential respite care. AIHW bulletin no. 78. Cat. no. AUS 124. Canberra: AIHW.
- AIHW 2011a. Pathways in aged care: do people follow recommendations? AIHW bulletin no. 88. Cat. no. AUS 137. Canberra: AIHW.
- AIHW 2011b. Pathways in Aged Care: program use after assessment. Data linkage series no. 10. Cat. no. CSI 10. Canberra: AIHW.
- AIHW 2013. Movement between hospital and residential aged care 2008–09. Data linkage series no. 16. Cat. no. CSI 16. Canberra: AIHW.
- AIHW 2014. Patterns in use of aged care: 2002–03 to 2010–11. Data linkage series no. 18. Cat. no. CSI 20. Canberra: AIHW.
- AIHW 2015. Use of aged care services before death. Data linkage series no. 19. Cat. no. CSI 21. Canberra: AIHW.
- AIHW: Karmel R 2005a. Data linkage protocols using a statistical linkage key. Data linkage series no. 1. Cat. no. CSI 1. Canberra: AIHW.
- AIHW: Karmel R 2005b. Transitions between aged care services. Data linkage series no. 2. Cat. no. CSI 2. Canberra: AIHW.
- AIHW: Karmel R & Braun P 2004. Statistical linkage across aged care programs: an exploratory example. Welfare working paper no. 44. Canberra: AIHW.
- AIHW: Ryan T, Holmes B & Gibson D 1999. A National Minimum Data Set for Home and Community Care. Aged care series. Cat. no. AGE 13. Canberra: AIHW.
- Guiver T 2011. Sampling-Based Clerical Review Methods in Probabilistic Linking. ABS Cat. no. 1351.0.55.034. Canberra: ABS.
- Karmel R, Anderson P, Gibson D, Peut A, Duckett SJ & Wells Y 2010. Empirical aspects of record linkage across multiple data sets using statistical linkage keys: the experience of the PIAC cohort study. BMC Health Services Research 10/41.
- Karmel R, Gibson D, Anderson P, Wells Y & Duckett S 2012. Care trajectories through community and residential aged care services: disease effects. Ageing and Society 32: 1428–45.


List of tables

- Table 2.1: Overview of PIAC data sources 5
- Table 5.1: ‘Master’ demographic variables for PIAC 2014 23
- Table 5.2: Service use variables for PIAC 2014 24
- Table 5.3: Characters used in variables event/event2 for PIAC 2014 25
- Table 5.4: ACAP MDS variables for PIAC 2014 26
- Table 5.5: Pathways variables for PIAC 2014 27
- Table 5.6: Characters used in variables straight_path/simple_path/path_repeats/path_skel/
path_care/path_care_changes for PIAC 2014..... 28
- Table 5.7: Examples showing the relationship between the different types of variables
available for ‘pathways’ 29

Related publications

This document, *Pathways in Aged Care 2014: Technical user guide*, is intended as an introductory manual for how to use the link map. Further information on the data linkage process is available in previous PIAC publications, particular *Patterns in use of aged care: 2002–03 to 2010–11*.

For detailed information on RCS and ACFI, please seek out the user guides applicable to the year in question. The NACDC metadata documentation also provides some information on these data items, as well as on all the tables contained within the NACDC data holdings. An overview of NACDC data is available in the *National Aged Care Data Clearinghouse Data Dictionary*.



This document is designed to support users of the Pathways in Age Care (PIAC) link map. It provides an introduction to the linkage process that created it, describing the various data sources that can be brought together by the link map. It also provides guidance on using the data.

aihw.gov.au



Stronger evidence,
better decisions,
improved health and welfare

