



Australian Government

Australian Institute of
Health and Welfare

Improving Indigenous identification in mortality estimates

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.

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 2019 

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.

ISBN 978-1-76054-551-2 (Online)

ISBN 978-1-76054-552-9 (Print)

Suggested citation

Australian Institute of Health and Welfare 2019. Improving Indigenous identification in mortality estimates. Cat. no. IHW 215. Canberra: AIHW.

Australian Institute of Health and Welfare

Board Chair

Mrs Louise Markus

Chief Executive Officer

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.

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 amendment.

Contents

Summary.....	ii
1 Introduction.....	1
2 Improving the collection of Indigenous status information	3
3 Misclassification of Indigenous status	4
4 The extent of Indigenous misclassification in death records	7
5 Enhancing Indigenous identification in mortality estimates	8
6 Denominator populations for preparing estimates	11
7 Conclusions.....	13
Appendix A: AIHW studies focusing on improving Indigenous identification in data.....	14
Acknowledgments	18
Abbreviations	18
Glossary	19
References	21
List of tables	23
List of figures.....	23

Summary

Measuring the extent of progress towards improvement in Indigenous health and welfare outcomes relies on consistent, complete, and reliable identification of Indigenous Australians in key data collections.

Indigenous identification is often incomplete, or is inconsistently reported across data sets. Incomplete and inconsistent reporting of Indigenous identification occurs through a combination of misclassification by service providers, and Indigenous people not identifying in certain circumstances. Incomplete and inconsistent reporting of Indigenous status usually results in the Indigenous population and their use of services being underestimated, which has an impact on the accurate planning and delivery of services to Aboriginal and Torres Strait Islander people.

To support Australian, and state and territory governments' Closing the Gap initiatives and improve reporting of outcomes for Indigenous Australians, the AIHW has, over the past 15 years, worked to assess the quality of Indigenous identification in key data sets, and to develop methods to adjust for under identification.

This report provides an overview of work undertaken by the AIHW to improve the estimation of Indigenous mortality and life expectancy measures, using statistical data linkage. This work was done under the Enhanced Mortality Database project, and covered the periods 2001–2005, 2006–2010 and 2011–2015.

It involved linking registered death records in the National Death Index (NDI) to selected data sets containing information on deaths and Indigenous identification. These data sets were hospital separation records, residential aged care separation records, and midwives or perinatal data collections.

Indigenous identification across the linked data sets was compared. The result was used to develop algorithms and adjustment factors to derive enhanced Indigenous identification for records with inconsistently reported Indigenous status, and those with 'not stated', 'don't know', and missing Indigenous status values. These records are deemed to be misclassified (see Chapter 3).

Results from the Enhanced Mortality Database project showed that during 2001–2015, about 13.6% of male and 13.9% of female Indigenous death records had been misclassified. These represent records that were not originally identified on official death records as Indigenous, but were deemed to be Indigenous through data linkage.

The misclassified records varied by age, sex, and state and territory of usual residence. So, mortality data were adjusted by age, sex, and state and territory of usual residence to produce the enhanced Indigenous mortality and life expectancy estimates.

Without the use of data linkage to enhance Indigenous status and make adjustments to the mortality estimates, Indigenous death rates would have been underestimated, and life expectancy estimates would have been overestimated by about 2.3 years for males, and about 2.1 years for females in 2011–2015. Because of the size of the non-Indigenous population however, the impact of Indigenous under-identification on non-Indigenous measures is often negligible.

The impact of the choice of denominator population on the mortality estimates was also investigated, using official back-cast Indigenous population estimates based on the 2011 Census, and cohort-interpolated Indigenous population estimates based on the 2001, 2006, 2011, and 2016 Censuses.

For the periods 2001–2005 and 2006–2010, life expectancy estimates based on the back-cast Indigenous population estimates were higher than those based on the cohort-interpolated Indigenous population. This is because denominator populations based on the back-cast Indigenous population estimates were larger than denominator populations based on the cohort-interpolated Indigenous population estimates.

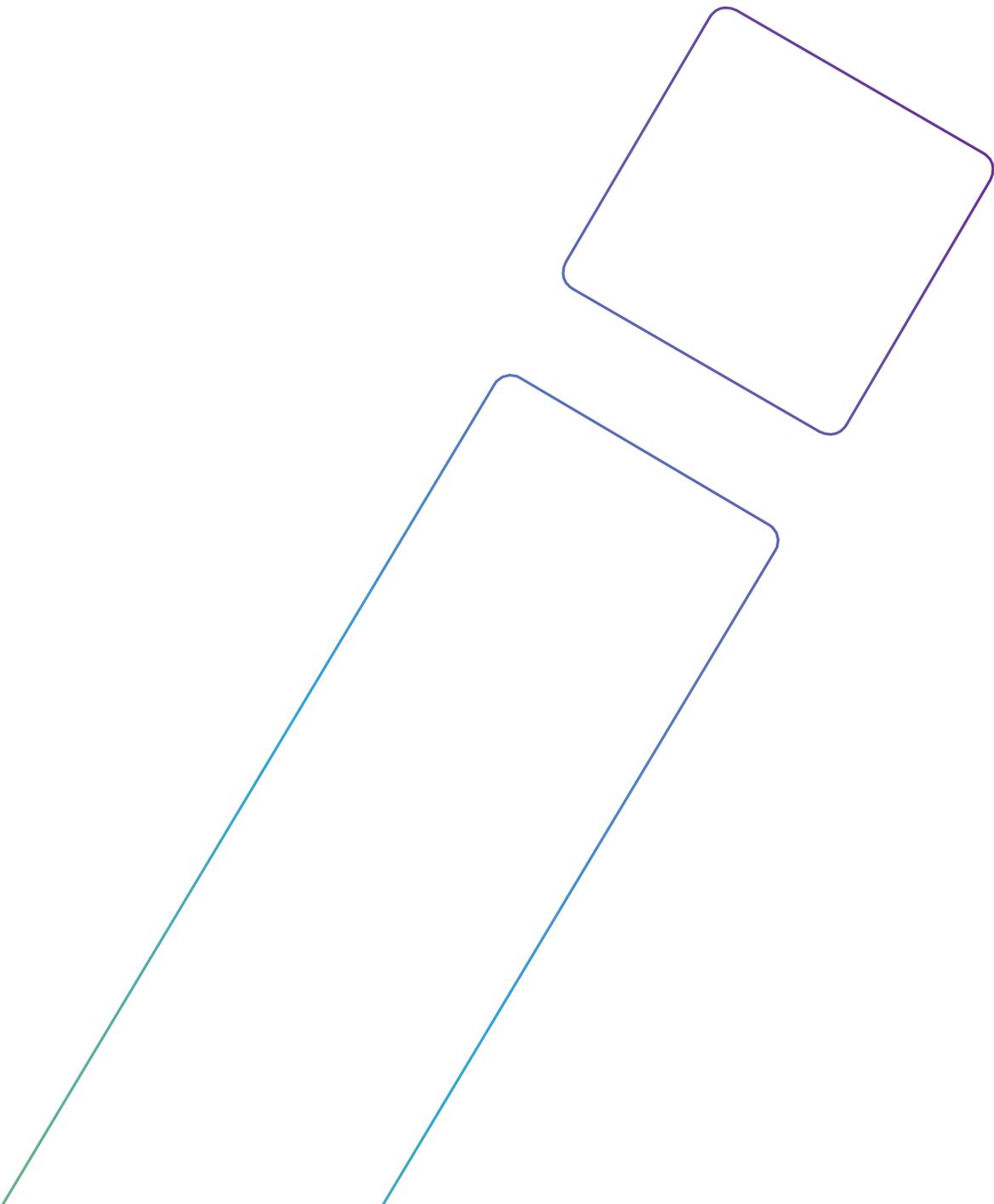
These larger denominator populations resulted in lower age-specific death rates, which, in turn, produced higher life expectancy estimates.

The AIHW proposes to test life expectancy estimates based on the back-cast of the 2016 Census when they become available.

Therefore, the various factors that could affect estimates of Indigenous mortality and life expectancy are:

- Indigenous identification on death records
- the method of enhancing Indigenous identification on death records, including age–sex adjustment
- the choice of a denominator population.

The impact of these factors should be considered when producing Indigenous mortality and life expectancy estimates.



1 Introduction

Improving outcomes for Aboriginal and Torres Strait Islander people is a high priority for all levels of government. In 2008, the Council of Australian Governments (COAG) committed to a set of targets to halve or eliminate the gap between Indigenous and non-Indigenous Australians in several areas (COAG 2008; PM&C 2015). Targets were set for:

- life expectancy
- child mortality
- early childhood education
- reading
- writing and numeracy
- Year 12 attainment
- employment outcomes
- school attendance (added in 2014).

In December 2018, COAG committed to work with Indigenous Australians on a refreshed closing the gap agenda (COAG 2018).

Complete, reliable, and consistently-reported Indigenous identification when accessing or registering for a service is important for policy formulation, program development, and service delivery. Currently, progress on some outcomes for Indigenous Australians is difficult to accurately measure where the targets rely on administrative data, as many Indigenous Australians may not be identified in all circumstances or are misclassified when accessing or registering for services.

The Australian Institute of Health and Welfare (AIHW) is working to highlight issues with Indigenous identification across health data sets, and to find ways to improve the quality of recording of Indigenous identification at the point of data collection. The AIHW is also developing guidelines and application of statistical methods to improve available data sets.

Appendix A includes a more extensive list of key AIHW work in this area, but largely this work includes:

- developing best-practice guidelines for collecting Indigenous status information in key health data sets (AIHW 2010a), along with online resources through the AIHW National Indigenous Data Improvement Support Centre, and online resources to help train staff on how to ask the question and collect the information

- assessing the quality of Indigenous identification across key data collection areas in the health sector, including:
 - perinatal data (AIHW 2007a)
 - pathology forms (AIHW 2013a)
 - primary health care data (AIHW 2013b, 2013c)
 - hospital admission records (AIHW 2005, 2010b, 2013d)
 - alcohol and other drug treatment services (AIHW 2004, 2006, 2007b)
 - mental health services (AIHW 2009)
 - community services data collections (AIHW 2012a)
- developing and publishing enhanced Indigenous estimates, using statistical methods, including data audits on hospital admission records (AIHW 2005, 2010b, 2013d) and data linkage using algorithms to measure and adjust for the extent of Indigenous under-identification in mortality data indicators (AIHW 2012b, 2017)
- collecting information on Indigenous babies, not just babies born to Indigenous mothers (AIHW 2007a, 2007c; AIHW: Leeds et al. 2007; AIHW: Li et al. 2013).

This report provides an overview of AIHW work in improving the estimation of Indigenous mortality and life expectancy measures, particularly through statistical data linkage.

2 Improving the collection of Indigenous status information

The starting point for more reliable reporting on Indigenous Australians is collecting reliable and consistent information on Indigenous identification across data sets at the point of service provision or the point where the information is being collected.

The AIHW has developed guidelines (AIHW 2010a) to help service providers across Australia collect more complete and reliable information on Indigenous identification across health data sets. The guidelines are also suitable for collecting Indigenous status information on other data sets.

The data collection guidelines outline that all clients should be asked the standard question on Indigenous identification to establish whether or not they are Aboriginal and/or Torres Strait Islander. The guidelines recommend that self-reporting is the most accurate means of ascertaining a person's Indigenous or non-Indigenous status. So, it is important to ask the standard question of all clients, irrespective of appearance, country of birth, or whether the staff know of the client and their family background. According to the guidelines, the standard response options must be provided to all clients to answer the question (either verbally or on a written form)—(see Box 2.1).

Box 2.1: Standard question on Indigenous identification

[Are you] [Is the person] [Is (name)] of Aboriginal or Torres Strait Islander origin?

- No
- Yes, Aboriginal
- Yes, Torres Strait Islander

For clients who identify as both Aboriginal and Torres Strait Islander, both 'Yes' boxes should be marked. Alternatively, a fourth response category, shown below, may also be included if this better suits the data collection practices of the agency or establishment concerned.

- Yes, both Aboriginal and Torres Strait Islander

Responses to the question on Indigenous identification can then be coded as set out in the National Health Data Dictionary (AIHW 2003) as follows:

- Aboriginal but not Torres Strait Islander origin
- Torres Strait Islander but not Aboriginal origin
- both Aboriginal and Torres Strait Islander origin
- neither Aboriginal nor Torres Strait Islander origin
- not stated/inadequately described.

3 Misclassification of Indigenous status

The Commonwealth definition of an Aboriginal or Torres Islander is a person of Aboriginal or Torres Islander descent who identifies as an Aboriginal or Torres Strait Islander, and is accepted as such by the community in which he or she lives (ABS 1999).

The best-practice guidelines for collecting Indigenous status information stipulate that the standard question, which includes the first 2 parts of the definition (origin and self identification), should be asked of everyone receiving a service (AIHW 2010a).

But in practice, not all service providers ask the standard question of everyone—they might ask the question in an inconsistent manner, or they might ascribe an Indigenous status based on the person's appearance. In addition, where the standard question is asked, not all Indigenous Australians may identify as Indigenous on all occasions.

Together, these issues can lead to an Indigenous person being recorded as 'non-Indigenous' or as having a 'not-stated' or 'missing' Indigenous status, which might result in Indigenous under-identification or lower numbers of Indigenous Australians being identified as using a service. A non-Indigenous person could also be recorded as Indigenous as a result of a recording or system's error. This could lead to Indigenous over-identification. Instances of Indigenous under-identification and/or over-identification is referred to as Indigenous misclassification.

The term 'misclassified' is therefore used in this report and in the Enhanced Mortality Database project to refer to death records that were not originally classified as 'Indigenous' on the NDI but were deemed to be Indigenous after comparing the Indigenous status of these records with the Indigenous status of corresponding records on the linked data sets. As Indigenous identification is determined through self-identification, the term 'misclassified' is not intended to convey the impression that individuals mis-identified themselves or that the Indigenous classification on the NDI was incorrectly reported.

While Indigenous misclassification includes both under-identification and over-identification, this report focusses on the under identification component of Indigenous misclassification.

The propensity to identify or to be identified as Indigenous varies, depending on age, sex, where the person lives (jurisdiction of usual residence or remoteness area), the service being accessed, and over time. The Indigenous status of individuals is often different across data sets, and data linkage can be used by statistical agencies and researchers to enhance Indigenous identification.

But there are no nationally agreed approaches on how to determine Indigenous status for individuals where classification is inconsistent or missing across multiple data sets. This might result in different methods being used to determine Indigenous status, and difficulties in interpreting findings, particularly when comparing results across studies.

Indigenous misclassification affects the accurate estimation of Indigenous mortality and life expectancy. Although all deaths in Australia must be registered by law, not all Indigenous deaths are identified as Indigenous. Responses obtained to the Indigenous question during death registration might be wrongly recorded as 'non-Indigenous', 'don't know', or 'not stated'; the Indigenous status information might also be missing.

The consequence of this misclassification is that Indigenous mortality and life expectancy estimates are likely to be underestimated. For more details on varying approaches to the classification of records with non-stated, unknown, or missing responses, see Box 3.1.

The misclassification of Indigenous status on death records was explored by the AIHW in the Enhanced Mortality Database project, which looked at the feasibility of using data linkage to both measure the extent of misclassification in mortality data, and to enhance Indigenous identification in mortality estimates.

This project involved linking death records in the NDI to deaths recorded across hospital separation records, residential aged care records, and perinatal data. Once these data sets are linked, the Indigenous status of corresponding records across the linked data sets is used to enhance Indigenous identification in the NDI, and the death data can be used to produce more accurate estimates of Indigenous mortality and life expectancy (AIHW 2012b).

Box 3.1: Treatment of 'not stated', 'don't know', missing values, or blank responses

An important source of Indigenous under-identification is the number of records whose Indigenous status is classified as 'not stated', 'don't know', or 'missing'. The way these records are treated affects any resulting estimates. The following ways are used by various analysts:

- The records with the 'not stated', 'don't know', and missing values are excluded from calculating the measure of interest. This means only the reported Indigenous and non-Indigenous values are used to calculate the measure of interest. This approach can understate the true value of the measure.
- The 'not stated', 'don't know', and missing values among the 'Indigenous' and 'non-Indigenous' categories are pro-rated, depending on the proportional representation of the 'Indigenous' and 'non-Indigenous' categories in the distribution. This results in under-identification of Indigenous rates if the 'not stated', 'don't know', and missing values are more likely to be Indigenous than not.
- The data are categorised into 'Indigenous' and 'other', with 'other' comprising the non-Indigenous and 'not stated' or missing value responses together. This results in under-identification of Indigenous rates if most of the 'not stated', 'don't know', and missing values are Indigenous.
- The Indigenous status for 'not stated', 'don't know', and missing values are imputed. This will depend on having sufficiently reliable values on these and other records to enable Indigenous status to be reliably imputed.

Two reports have been published from the Enhanced Mortality Database project:

- The first used data linkage to measure the extent of Indigenous misclassification in mortality data, to develop algorithms that better identify Indigenous status in mortality data, and to improve estimates of Indigenous mortality and life expectancy in 2001–2005 (AIHW 2012b).
- The second, which had the same objectives, provided estimates of Indigenous mortality and life expectancy for 2001–2005, 2006–2010, and 2011–2015, using an improved method (AIHW 2017).

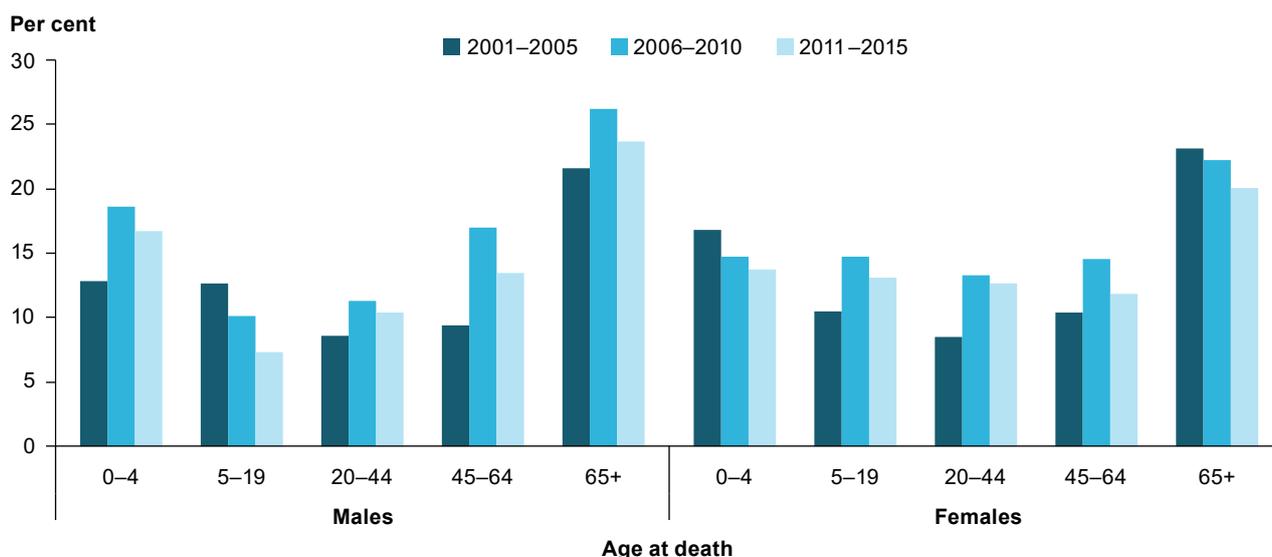
For the purpose of this study, the ever-Indigenous algorithm was used to enhance Indigenous identification on the death records. This means that if a death is identified as Indigenous in any of the data sets, it will be recorded as Indigenous in the NDI to estimate Indigenous mortality and life expectancy.

4 The extent of Indigenous misclassification in death records

Analysis of the Enhanced Mortality Database showed that:

- in 2001–2005, 783 male and 685 female Indigenous death registrations on the NDI had not been identified as Indigenous—about 12.9% of male and 15.3% of female death registrations originally identified as Indigenous on the NDI
- in 2006–2010, 1,208 male and 962 female Indigenous death registrations had not been identified as Indigenous—about 15.3% of male and 15.0% of female death registrations originally identified as Indigenous on the NDI
- in 2011–2015, 1,288 male and 986 female Indigenous death registrations had not been identified as Indigenous—about 15.9% of male and 15.5% of female death registrations originally identified as Indigenous in the NDI in 2011–2015 (Figure 4.1).

Figure 4.1: Misclassified Indigenous deaths as a percentage of records originally identified as Indigenous on the NDI, by sex and age group, 2001–2005, 2006–2010, 2011–2015



Source: Enhanced Mortality Database.

For both males and females, in all periods, the level of Indigenous misclassification in the NDI was highest in the youngest (0–4) and oldest (65 and over) age groups. There was very little change in either the level or age–sex pattern of Indigenous misclassification during the reference periods, although misclassification was slightly higher for both sexes in 2006–2010 than in the other 2 periods.

The levels, patterns, and trends of Indigenous mortality and life expectancy would have been incorrectly reported if they had been based on unadjusted death registrations.

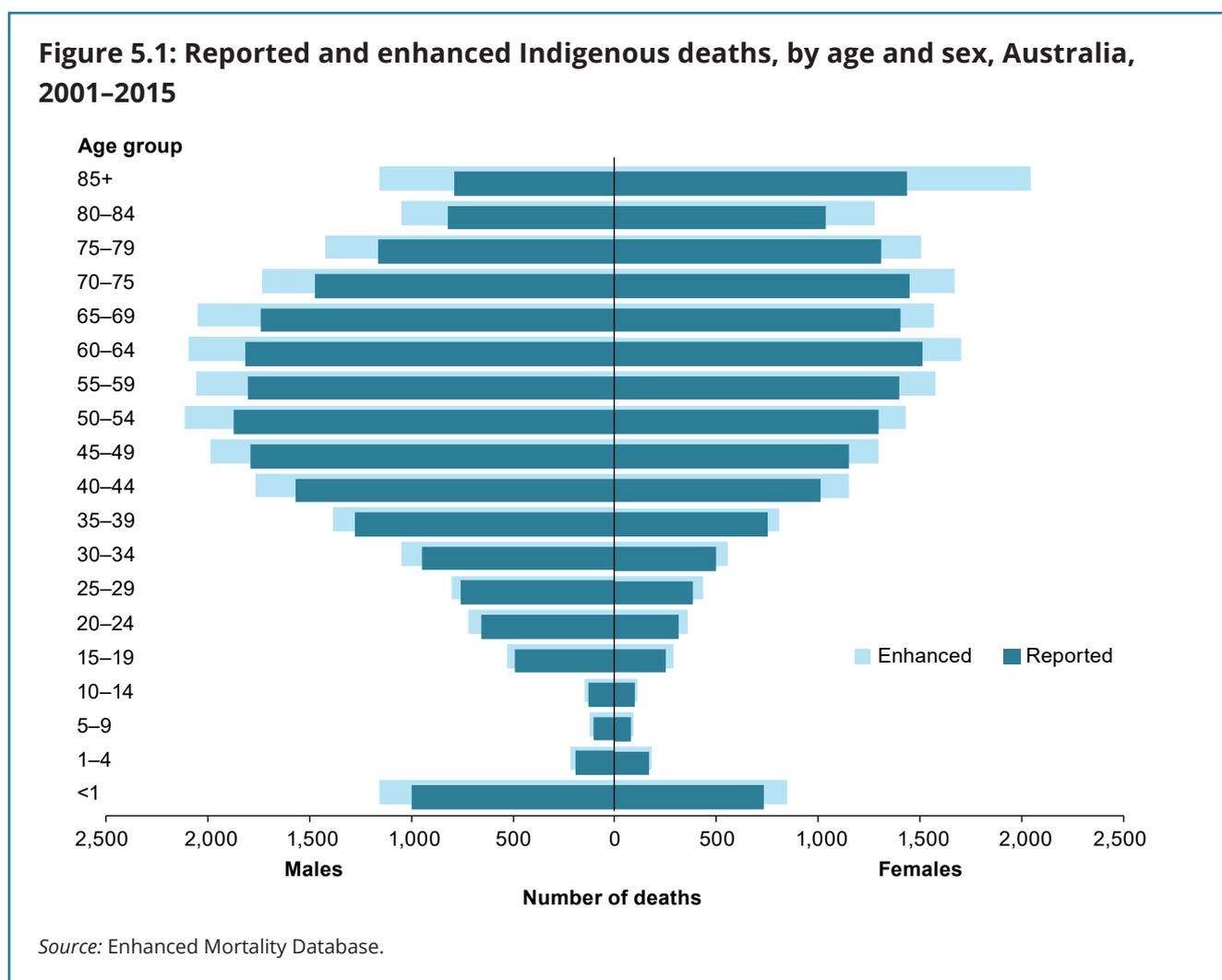
Because of differences in the age structure between the Indigenous and non-Indigenous populations, age-standardisation is used to enable more reliable comparison between Indigenous and non-Indigenous Australians on key measures and indicators, given most health measures and indicators are related to age (AIHW 2011).

5 Enhancing Indigenous identification in mortality estimates

Analysis of the Enhanced Mortality Database showed that 3,209 Indigenous male and 2,633 Indigenous female deaths in 2001–2015 had been misclassified in the NDI as ‘non-Indigenous’, or as records with ‘not stated’, ‘don’t know’, or ‘missing’ Indigenous status.

These records were subsequently reclassified as ‘Indigenous’, and added to the number of death records originally classified as ‘Indigenous’ in the NDI.

Figure 5.1 shows the distribution of reported and enhanced Indigenous deaths in Australia in 2001–2015. Full details of the reclassified records by age and sex during the 5-year periods 2001–2005, 2006–2010, and 2011–2015 can be found in the second Enhanced Mortality Database report (AIHW 2017).



Nationally, in 2001–2015, reclassified death records made up 13.6% of Indigenous male records and 13.9% of Indigenous female records (Table 5.1). There were slight variations between males and females in the percentage of reclassified death records during the individual reference periods.

In 2001–2015, the percentage of reclassified records was highest for the oldest (65 and over) age group. Although the number of deaths at the youngest (0–4) age group was relatively low, a large percentage had not been identified as Indigenous, and had to be reclassified. The lowest proportions of reclassified records were for those aged 5–19 and 20–44.

The number of reclassified Indigenous death registrations did not vary much between the reference periods, at about 11.4%–15.3% for males, and 13.3%–15.0% for females.

The proportion of Indigenous deaths that had to be reclassified varied considerably between the states and territories. Further details can be found in *Trends in Indigenous mortality and life expectancy 2001–2015* (AIHW 2017).

Mortality data were adjusted for the misclassified records for each age group and by sex within each state and territory, and nationally. The ever-Indigenous method was used to enhance the mortality data. These data were used to create life tables, and to estimate life expectancy over a 5-year period. The resulting life expectancy estimates have much lower sample errors, as they are based on 60 months of deaths data (AIHW 2017).

Table 5.1: Reclassified Indigenous deaths, by sex and age, 2001–2015 (%)

Age group	Males (%)	Females (%)
0–4	13.8	13.2
5–19	9.0	12.5
20–44	9.2	10.4
45–64	11.9	11.0
65+	19.3	17.7
All ages	13.6	13.9

Source: Enhanced Mortality Database.

When reclassified Indigenous deaths are excluded from the calculation of Indigenous life expectancy estimates, estimates of Indigenous life expectancy at birth are inflated (see Table 5.2).

For Indigenous males, life expectancy is inflated by:

- 2.6 years in 2001–2005
- 2.1 years in 2006–2010
- 2.3 years in 2011–2015.

For Indigenous females, life expectancy is inflated by:

- 1.6 years in 2001–2005
- 1.5 years in 2006–2010
- 2.1 years in 2011–2015 (Table 5.2).

Table 5.2: Indigenous life expectancy at birth, adjusted and not adjusted for misclassification, by sex, 2001–2005, 2006–2010, 2011–2015

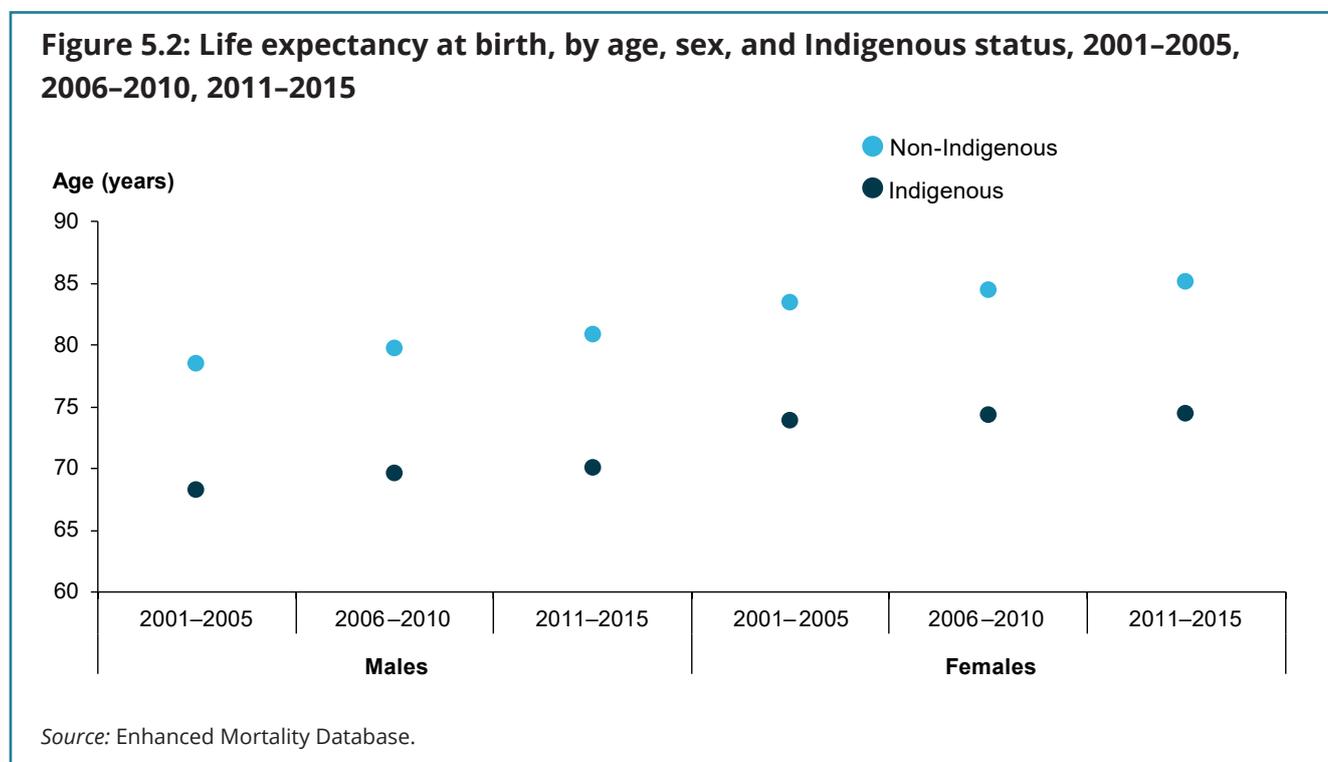
Australia	Males		Females	
	Adjusted for Indigenous misclassification	Not adjusted for Indigenous misclassification	Adjusted for Indigenous misclassification	Not adjusted for Indigenous misclassification
2001–2005	68.3	70.9	73.9	75.5
2006–2010	69.6	71.7	74.4	75.9
2011–2015 ^(a)	70.1	72.4	74.5	76.5

(a) Estimates are based on reported deaths for 2011–2015, which were enhanced using Indigenous reclassification rates for 2006–2010, based on linkage of the NDI with hospital, aged care, and perinatal data sets.

Source: Enhanced Mortality Database.

The difference in life expectancy at birth between estimates based on the reported and enhanced Indigenous deaths shows that misclassification has a big impact on estimates.

Figure 5.2 compares estimates of life expectancy at birth for Indigenous and non-Indigenous males and females in 2001–2005, 2006–2010 and 2011–2015. It shows that life expectancy at birth had increased for all groups, but it increased faster for non-Indigenous males and females than for Indigenous males and females. As a result, there was no narrowing of the gap in life expectancy between Indigenous and non-Indigenous Australians.



6 Denominator populations for preparing estimates

The reliability of rates, such as death rates, can be significantly at risk if the numerator (the number of deaths) and the denominator (the total population at risk of dying) are drawn from 2 different data sets, which might have been collected by different agencies, under different collection regimes, for different purposes, and subject to different levels of quality on most data items, including Indigenous identification.

In estimating death rates, for instance, the numerator, comprises the number of deaths in a defined period, compiled by jurisdictional Registries of Births, Deaths, and Marriages.

On the other hand, the denominator, or the population at risk, is a synthetic estimate of the population in the middle of the defined period, based on information collected during the Census of Population and Housing.

The 2 data sets are subject to different levels of accuracy of Indigenous identification, which will affect estimates of mortality.

The problem of numerator–denominator bias also exists in other rates, such as marriage and divorce rates, school participation rates, and hospital use rates, where:

- the numerator is based on administrative data on the number of people participating in the activity who are recorded as being Indigenous
- the denominator is a census-based estimate of the number of people in the middle of the year who identified as being Indigenous at the time of the census.

Indigenous misclassification is not the only issue that can affect the quality of Indigenous health and welfare statistics. Also important is the choice of the denominator population used in preparing Indigenous measures, such as death rates and life expectancy estimates.

In *Trends in Indigenous mortality and life expectancy 2001–2015* (AIHW 2017), the AIHW explored the use of 2 denominator populations for preparing Indigenous measures. These were:

- the official ABS back-cast and projected Indigenous population estimates, based on the 2011 Census and back-casted to 2001
- cohort-interpolated population estimates based on the 2001, 2006, and 2011 censuses.

The ABS back-cast and projected populations are based on a single census, while cohort-interpolated populations are based on multiple censuses.

Cohort-interpolated measures that cover 2001–2006 are based on denominator populations estimated from the 2001 and 2006 censuses, while measures covering 2006–2011 are based on denominator populations estimated from the 2006 and 2011 censuses.

There are strengths and weaknesses for the use of either of these 2 populations as denominators. The advantage of the back-cast and projected Indigenous populations is that they represent a demographically-consistent population. The ABS regards them as their most accurate and up-to-date estimate, and advises that, wherever possible, they should be used, but notes that there is a range of circumstances where an alternative denominator may be more appropriate (ABS 2014).

A key advantage of the approach using the cohort-interpolated population, however, is that the Indigenous identification in the denominator is likely to be aligned more closely with the Indigenous identification of the events in the numerator, (in this case deaths).

Table 6.1 compares estimates of Indigenous life expectancy at birth, based on the two denominator populations.

Table 6.1: Indigenous life expectancy at birth based on 2 denominator populations, by sex, 2001–2005, 2006–2010, 2011–2015

Years	Males		Females	
	Back-cast	Cohort-interpolated	Back-cast	Cohort-interpolated
2001–2005	68.3 (67.9–68.7)	65.9 (65.5–66.3)	73.9 (73.6–74.2)	72.2 (71.8–72.6)
2006–2010	69.6 (69.3–70.0)	68.2 (67.9–68.6)	74.4 (74.0–74.8)	73.3 (73.0–73.7)
2011–2015	70.1 (69.7–70.3)	70.5 (70.2–70.9) ^(a)	74.5 (74.2–74.8)	75.4 (77.0–75.7) ^(a)

(a) Provisional estimate not contained in (AIHW 2017).

Source: Table C7, AIHW 2017.

A comparison of estimates of Indigenous life expectancy at birth using the 2 denominator populations shows that:

- in 2001–2005 and 2006–2010, estimates based on the back-cast population are higher than estimates based on the cohort-interpolated population. This is because the back-cast denominator population based solely on the 2011 Census is larger than denominators derived by interpolating between the 2001, 2006, and 2011 censuses
- over 2011–2015, however, estimates of Indigenous life expectancy at birth based on the cohort-interpolated population are higher than those based on the back-cast population. This is because the cohort-interpolated denominator population based on the 2011 and 2016 Censuses is larger than the back-cast population, based solely on the 2011 Census (see Table 6.2).

Table 6.2: Estimates of Indigenous population based on the Census, by sex, at 30 June 2001, 2006, 2011, 2016

Date of population estimate	Males	Females
30 June 2001 based on 2001 Census	227,526	230,994
30 June 2006 based on 2006 Census	257,309	259,734
30 June 2011 based on 2011 Census	333,683	336,198
30 June 2016 based on 2016 Census	398,413	399,952

Sources: (ABS 2004, 2009, 2014, 2018)

Life expectancy estimates based on either the back-cast or cohort-interpolated denominator population do not involve a change to the numerator (deaths), but only a change to the denominator population. So the level of any resulting estimates will depend on the denominator population used in preparing the estimates. This means a larger denominator population will lead to lower age-specific death rates and higher life expectancy estimates (AIHW 2017).

7 Conclusions

The AIHW has undertaken a range of initiatives to improve the quality of Indigenous identification in key data sets, including improving the estimation and reporting of Indigenous mortality and life expectancy measures.

These initiatives include methods aimed at improving the collection of more consistent, complete and reliable Indigenous identification in health data sets at the point of data collection, and statistical methods, including data linkage and data audits, to assess and adjust for Indigenous under-identification on death records and hospital admission data.

The AIHW is in the process of finalising data access to produce an ongoing enhanced Indigenous mortality data collection (EIMDC). The EIMDC will be used to prepare ongoing statistics of mortality and life expectancy.

Once the EIMDC is established and ongoing, the AIHW will undertake the following activities:

- Include cause of death on the EIMDC to aid investigation of the contribution of cause of death to the levels, trends and patterns of Indigenous mortality as well as to the gap in mortality and life expectancy between Indigenous and non-Indigenous Australians.
- Explore methods to estimate life expectancy for jurisdictions with small Indigenous populations for which official life expectancy estimates are currently not available (Victoria, South Australia, Tasmania, and the Australian Capital Territory), as well as for lower levels of geography.
- Continue to explore the impact of using alternative denominator populations, including cohort-interpolated denominator populations, for the estimation of Indigenous mortality and life expectancy measures.
- Expand the number and scope of relevant data sets used in creating the EIMDC in order to improve the quality of linkage and the estimation of Indigenous mortality and life expectancy measures.
- Explore a range of algorithms that can be used to enhance Indigenous identification on mortality data beyond the “ever-Indigenous” method used in current estimates.
 - the wider the choice of algorithms that can be used in identifying Indigenous misclassification on death data, the more reliable will be the mortality and life expectancy estimates based on the reclassified death data.

Appendix A: AIHW studies focusing on improving Indigenous identification in data

Table A1.1: Key studies towards improving Indigenous identification in data

Name of study/report	Link	Description
Methods to improve the collection of Indigenous status in data		
<i>National best practice guidelines for collecting Indigenous status in health data sets</i> (AIHW 2010a)	https://www.aihw.gov.au/getmedia/ad54c4a7-4e03-4604-a0f3-ccb13c6d4260/11052.pdf.aspx?inline=true	Despite improvements in recent years, there have been continuing problems in establishing and maintaining standard practice in the collection of Indigenous status, resulting in the under-identification of Aboriginal and Torres Strait Islander people in key national health data sets. These guidelines aim to ensure the standard Indigenous status question is asked correctly and consistently of all clients of health services, and that this information is correctly recorded.
Online resources to help train staff to collect Indigenous status in data collections		
<i>Staff knowledge training tool for Indigenous identification: version 3</i> (AIHW 2019)	https://360.articulate.com/review/content/ef93b47f-ddb5-48d2-af65-f17129c43a94/review	This is a training tool and companion document to the <i>National best practice guidelines for collecting Indigenous status in health data sets</i> (AIHW 2010a).
<i>General practice software model to support best practice in Indigenous health care</i> (AIHW 2015)	https://www.aihw.gov.au/getmedia/1e7b5110-3b61-445a-8f66-10ea1fb7305d/AIHW-GP-software-model.pdf.aspx	This is a software and tool to help general practitioners (GPs) improve the collection of Indigenous status information during clinical consultations. The software also provides information on Aboriginal and Torres Strait Islander-specific health measures that can be accessed via GPs. Companion documents are (AIHW 2013b, 2013c).
Assessment of the quality of Indigenous identification in various data sets		
<i>The inclusion of Indigenous status on pathology request forms</i> (AIHW 2013a)	https://www.aihw.gov.au/getmedia/33d85a0d-6e57-47c5-b9e7-bba313dcfc5c/15247.pdf.aspx?inline=true	Under the National Indigenous Reform Agreement in 2008, COAG agreed to data quality improvements that are focused on improving Indigenous identification in key data sets. This report outlines work towards including Indigenous status on pathology request forms as a way to improve Indigenous identification in national cancer, communicable disease, and cervical screening registries.
Assessment of Indigenous identification in primary health care data collection		This report will help the general practice sector collect the Indigenous status of patients to provide GP-mediated health interventions specific to Aboriginal and Torres Strait Islander people, and to provide Indigenous status data to national data collections. Effective Indigenous identification processes are not widespread in the sector, and this undermines both service delivery and data collection. This report describes the problem and discusses how Indigenous identification could be improved.
<i>Taking the next steps: identification of Aboriginal and Torres Strait Islander status in general practice</i> (AIHW 2013b)	https://www.aihw.gov.au/getmedia/6c8e19b9-2f9d-4ede-8e1f-bd63409ab1b7/15599.pdf.aspx?inline=true	

(continued)

Table A1.1 (continued): Key studies towards improving Indigenous identification in data

Name of study/report	Link	Description
Assessment of Indigenous identification in the mental health-care sector <i>Towards better Indigenous health data</i> (AIHW 2013c)	https://www.aihw.gov.au/getmedia/b424d472-2e23-4624-a669-22c7e1180036/15020.pdf.aspx?inline=true	The collection of the Indigenous status of patients/clients by health-care providers is important for improving Aboriginal and Torres Strait Islander health. This report provides recommendations on how to improve the collection of Indigenous status in hospital admission records, perinatal data collections, the drug and alcohol treatment services sector, the mental health sector, the National Diabetes Register, cancer registries, and general practice.
Assessment of Indigenous identification in community services data collections		
<i>Data quality of Aboriginal and Torres Strait Islander identification: seven community services data collections</i> (AIHW 2004)	https://www.aihw.gov.au/getmedia/e30e5405-0077-432f-b267-24fca403c54a/dqatsii.pdf.aspx?inline=true	This report looks at the quality of identification of Indigenous clients in 7 data collections. The analyses focused on: <ul style="list-style-type: none"> • the extent to which Indigenous status was missing or not stated in each data collection • where possible, whether improvements in data quality had occurred over time • variations in data quality in relation to particular groups of clients (for example, age, sex, or service type), and by geographic location.
<i>Quality of Aboriginal and Torres Strait Islander identification in community services data collections: update on eight community services data collections</i> (AIHW 2007b)	https://www.aihw.gov.au/getmedia/122edbdba-b05a-4224-acde-a45e2b4bdd9b/qatsiicsdc.pdf.aspx?inline=true	This report looks at the quality of identification of Indigenous clients in 8 community services data collections, by analysing the extent to which Indigenous status is missing or not stated in each of the data collections. The report highlights changes in Indigenous identification rates, and documents data quality improvement activities in each of the community services data collections, both at the national and the jurisdictional level. This report is useful for administrators of programs, and researchers with an interest in Aboriginal and Torres Strait Islander clients in the community services sector and identification issue.
<i>Aboriginal and Torres Strait Islander identification in community services data collections: an updated data quality report</i> (AIHW 2012a)	https://www.aihw.gov.au/getmedia/5a206296-eefa-4a43-b4be-e42673bf81c8/14677.pdf.aspx?inline=true	This report looks at the identification of Indigenous clients in various AIHW community services data collections, by analysing where Indigenous status is missing or not stated. It makes several recommendations, including that data collection manuals and training materials reflect the <i>National best practice guidelines for collecting Indigenous status in health data sets</i> . Where necessary, jurisdictions should consider modifying client forms and client information management systems to ensure consistency with these guidelines.

(continued)

Table A1.1 (continued): Key studies towards improving Indigenous identification in data

Name of study/report	Link	Description
Assessment of the level of Indigenous identification and enhancement of Indigenous identification in mortality data		
<i>National best practice guidelines for data linkage activities relating to Aboriginal and Torres Strait Islander people</i> (AIHW & ABS 2012)	https://www.aihw.gov.au/getmedia/6d6b9365-9cc7-41ee-873f-13e69e038337/13627.pdf.aspx?inline=true	To ensure a consistent approach to dealing with Indigenous status reporting, COAG requested that the AIHW and ABS develop national best-practice guidelines for linking data related to Indigenous Australians. This report offers guidance on how to derive Indigenous status when it is missing or inconsistently reported, in a way that protects privacy, and people's right to self-identification.
<i>An enhanced mortality database for estimating Indigenous life expectancy: a feasibility study</i> (AIHW 2012b)	https://www.aihw.gov.au/getmedia/3ae4bae6-0777-4faf-acd2-7f3c6e25fe8c/12664.pdf.aspx?inline=true	This study aims to improve reporting of information on Indigenous deaths, by linking death registrations data with several additional data sets that contain information on Indigenous deaths, and comparing information on Indigenous identification for corresponding records across the linked data sets. The result of the comparison is used to develop algorithms to enhance the Indigenous identification on the death records. The enhanced death records are used to prepare more robust estimates of Indigenous mortality and life expectancy.
<i>Trends in Indigenous mortality and life expectancy 2001–2015: evidence from the Enhanced Mortality Database</i> (AIHW 2017)	https://www.aihw.gov.au/getmedia/bbe476f3-a630-4a73-b79f-712aba55d643/aihw-ihw-174.pdf.aspx?inline=true	This report looks at Indigenous mortality and life expectancy in 2001–2015, based on evidence from the Enhanced Mortality Database. The study observed increases in life expectancy during the study period for both Indigenous males and females across most jurisdictions. But the gap between Indigenous and non-Indigenous Australians did not narrow, because life expectancy increased faster for non-Indigenous males and females than for Indigenous males and females.
Data audits		
<i>Improving the quality of Indigenous identification in hospital separations data</i> (AIHW 2005)	https://www.aihw.gov.au/getmedia/bde2cc69-c713-4da3-a95b-d902a4f9125d/iqiihs.pdf.aspx?inline=true	This report presents the outcomes of a project funded by the Australian Health Ministers' Advisory Council and directed at establishing a basis for improving Indigenous identification in hospital separations data.
<i>Indigenous identification in hospital separations data: quality report</i> (AIHW 2010b)	https://www.aihw.gov.au/getmedia/f0d647fa-61a4-4f02-81c3-f05c433815d6/10637.pdf.aspx?inline=true	This report presents the results of studies done in 2006–2008 by the AIHW, in collaboration with the state and territory health authorities, in a follow-up assessment of the quality of Indigenous identification in public hospitals. It presents revised recommendations for analysis of hospital separations data, based on the results of the studies, and estimates of correction factors that can be applied to the data for analysis purposes.

(continued)

Table A1.1 (continued): Key studies towards improving Indigenous identification in data

Name of study/report	Link	Description
<i>Indigenous identification in hospital separations data: quality report</i>	http://www.aihw.gov.au/publication-detail/?id=60129543215&tab=2	This report presents the results of a study on the quality of Indigenous identification in records of hospitalisations in public hospitals in Australia. It presents an analysis of hospital separations data, and estimates of correction factors that can be applied to the data for analysis purposes at 4 levels: national; national by remoteness area; state and territory; and remoteness area within jurisdictions. An estimated 88% of Indigenous patients were correctly identified in public hospital admission records in 2011–12. The report recommends that data for all jurisdictions be included in any analysis of hospitalisations by Indigenous status, and that correction factors be used to adjust total hospital data from 2010–11 onwards.
Statistical methods to enable comparison of differences and trends in Indigenous and non-Indigenous measures		
<i>Principles on the use of direct age-standardisation in administrative data collections: for measuring the gap between Indigenous and non-Indigenous Australians (AIHW 2011)</i>	https://www.aihw.gov.au/getmedia/95237794-4b77-4683-9f00-77c4d33e0e7c/13406.pdf.aspx?inline=true	This report recommends that the direct method of age standardisation be used to compare health and welfare outcome measures (for example, mortality rates, life expectancy, hospital separation rates, disease incidence rates etc.) of Indigenous and non-Indigenous Australians. The report provides consistency and guidance on when and how to use the direct age standardisation method, and when it should not be used.

Acknowledgments

Tetteh Dugbaza, Fadwa Al-Yaman, and Elizabeth Hynes prepared this report.

The authors are grateful to Len Smith (Australian National University), and Ching Choi (University of New South Wales) for useful comments on earlier drafts of the report.

Abbreviations

ABS	Australian Bureau of Statistics
AIHW	Australian Institute of Health and Welfare
CI	confidence interval
COAG	Council of Australian Governments
EIMDC	Enhanced Indigenous Mortality Data Collection
EMD	Enhanced Mortality Database
GP	general practitioner
NDI	National Death Index

Glossary

Aboriginal and Torres Strait Islander people: People who identify, or are identified, as being of Aboriginal and/or Torres Strait Islander origin. See also **Indigenous Australian**.

administrative data: Information that is required by law or that is collected for the purpose, or in the process, of service delivery, such as providing health care (National Hospital Morbidity Database), responding to the legal requirements of registering particular events (births and deaths registration data), or providing a particular service (residential aged care data set).

age-specific death rate: The number of deaths (registered or occurred) during the calendar year in a specified age or age group per unit of population, usually 1,000 or 100,000, of the same age or age group at the mid-point of the year.

age-standardisation: A method of removing the influence of age when comparing populations with different age structures. This is usually necessary because death rates vary strongly with age. The age structures of the different populations are converted to the same 'standard' structure, then the death rates that would have occurred with that structure are calculated and compared.

algorithm: A process or set of rules used for calculation or problem solving. In this report, 'algorithm' refers to a set of rules used to determine the Indigenous status of an individual based on a linked data set.

confidence interval (CI): A statistical term describing a range (interval) of values within which one can be confident that the true value lies, usually because it has a 95% or higher chance of doing so.

data linkage: The process of bringing together 2 or more sets of information belonging to the same person, event, or place into a single record of information. See also **linked record**.

enhanced Indigenous death records: Indigenous death records that comprise those that were originally identified as Indigenous on death registration data, and those that were not, but that were subsequently reclassified as Indigenous, after linkage with other data sets and comparison of their Indigenous status information across corresponding records on the linked data sets.

Indigenous identification: The process of identifying, or being identified, as being of Aboriginal and/or Torres Strait Islander origin.

Indigenous Australian: A person who identifies, or is identified, as being of Aboriginal and/or Torres Strait Islander origin. See also **Aboriginal and Torres Strait Islander people**.

Indigenous status: The name of the variable that describes whether or not a person identifies, or has been identified, as being of Aboriginal and/or Torres Strait Islander origin.

Indigenous under-identification: A situation that might occur if Indigenous status is not correctly collected and recorded for all clients. While this can also lead to over-identification, the tendency has often been for Aboriginal and Torres Strait Islander people to be recorded as being non-Indigenous, or for their Indigenous status not to be recorded at all.

life expectancy: An indication of how long a person can expect to live for, based on the age they have already reached. Technically, it is the number of years of life remaining to a person at a particular age if death rates do not change. The most commonly used example is life expectancy at birth.

life table: A representation of the number of years people in a defined population are likely to live for. It comprises any of various tables describing mortality and survival data for groups of individuals at specific times or over defined intervals. Life tables might summarise combined mortality experience by age over a brief period—usually 1–3 years (period life table)—or might follow a cohort over time (cohort life table).

linked record: A record that has passed through the data linkage process, and has been linked to a record from another file.

misclassified death: An Indigenous death record whose Indigenous status was recorded as ‘non-Indigenous’, ‘unknown’, or ‘not stated’ on death registration data, or an Indigenous death record whose Indigenous status on death registration data was missing.

numerator–denominator bias: A bias arising where the numerator and denominator of a rate or statistical measure are derived from different populations. This might occur when different data sources are used in the numerator and denominator, and where the data in the numerator and denominator are collected and/or compiled, often by different agencies, under different conditions, and for different purposes. An example is mortality rates, where the numerator is the number of deaths compiled by the registries of births, deaths, and marriages, while the denominator is the estimated resident population compiled from Census and other data.

reclassified Indigenous death record: An Indigenous death record that was not recorded as Indigenous on death registration data, but was deemed to be Indigenous, and subsequently reclassified as Indigenous, after being linked with comparative data sets, and after their Indigenous status information was compared with corresponding records across the linked data sets.

References

- ABS (Australian Bureau of Statistics) 1999. Standards for statistics on cultural and language diversity. Cat. no. 1289.0. Canberra: ABS.
- ABS 2004. Experimental estimates and projections, Aboriginal and Torres Strait Islander Australians, 30 June 1991 to 30 June 2009. ABS cat. no. 3238.0. Canberra: ABS.
- ABS 2009. Experimental estimates and projections, Aboriginal and Torres Strait Islander Australians. ABS cat. no. 3238.0. Canberra: ABS.
- ABS 2014. Estimates and projections, Aboriginal and Torres Strait Islander Australians, 2001 to 2026. ABS cat. no. 3238.0. Canberra: ABS.
- ABS 2018. Estimates of Aboriginal and Torres Strait Islander Australians, June 2016. Cat no. 3238.0.55.001. Canberra: ABS.
- AIHW (Australian Institute of Health and Welfare) 2003. National health data dictionary, version 12. Volumes 1 and 2. Cat. no. HWI 43 and HWI 44. Canberra: AIHW.
- AIHW 2004. Data quality of Aboriginal and Torres Strait Islander identification: Seven community services data collections. Cat. no. HWI 79. Canberra: AIHW.
- AIHW 2005. Improving the quality of Indigenous identification in hospital separations data. Cat. no. HSE 101. Health services series no. 25. Canberra: AIHW.
- AIHW 2006. Drug use among Aboriginal and Torres Strait Islander peoples: an assessment of data sources. Cat. no. PHE 76. Canberra: AIHW.
- AIHW 2007a. Australia's mothers and their babies, Australia 2001–2004. Cat. no. PER 38. Canberra: AIHW.
- AIHW 2007b. Quality of Aboriginal and Torres Strait Islander identification in community services data collections: update on eight community services data collections. Cat. no. HWI 95. Canberra: AIHW.
- AIHW 2007c. Australia's mothers and babies 2005. Cat. no. PER 40. Canberra: AIHW.
- AIHW 2009. Measuring the social and emotional wellbeing of Aboriginal and Torres Strait Islander peoples. Cat. no. IHW 24. Canberra: AIHW.
- AIHW 2010a. National best practice guidelines for collecting Indigenous status in health data sets. Cat. no. IHW 29. Canberra: AIHW.
- AIHW 2010b. Indigenous identification in hospital separations data: quality report. Cat. no. HSE 85. Canberra: AIHW.
- AIHW 2011. Principles on the use of direct age-standardisation in administrative data collections: for measuring the gap between Indigenous and non-Indigenous Australians. Cat. no. CSI 12. Canberra: AIHW.
- AIHW 2012a. Aboriginal and Torres Strait Islander identification in community services data collections: an updated data quality report. Cat. no. IHW 80. Canberra: AIHW.
- AIHW 2012b. An enhanced mortality database for estimating Indigenous life expectancy: a feasibility study. Cat. no. IHW 75. Canberra: AIHW.
- AIHW 2013a. The inclusion of Indigenous status on pathology request forms. Cat. no. IHW 103. Canberra: AIHW.

AIHW 2013b. Taking the next steps: identification of Aboriginal and Torres Strait Islander status in general practice. Cat. no. IHW 100. Canberra: AIHW.

AIHW 2013c. Towards better Indigenous health data. Cat. no. IHW 93. Canberra: AIHW.

AIHW 2013d. Indigenous identification in hospital separations data: quality report. Cat. no. IHW 90. Canberra: AIHW.

AIHW 2015. General practice software model to support best practice in Indigenous health care. Canberra: AIHW. Viewed 3 May 2019. <https://www.aihw.gov.au/getmedia/1e7b5110-3b61-445a-8f66-10ea1fb7305d/AIHW-GP-software-model.pdf.aspx>

AIHW 2017. Trends in Indigenous mortality and life expectancy 2001–2015: evidence from the Enhanced Mortality Database. Cat. no. IHW 174. Canberra: AIHW.

AIHW 2019. Staff knowledge training tool for Indigenous identification. Version 3. Canberra: AIHW. Viewed 3 May 2019. <https://360.articulate.com/review/content/ef93b47f-ddb5-48d2-af65-f17129c43a94/review>

AIHW & ABS 2012. National best practice guidelines for data linkage activities relating to Aboriginal and Torres Strait Islander people. Canberra: AIHW.

AIHW: Leeds K, Gourley M, Laws P, Zhang J, Al-Yaman F & Sullivan E 2007. Indigenous mothers and their babies, Australia 2001–2004. Perinatal statistics series no. 19. Cat. no. PER 38. Canberra: AIHW.

AIHW: Li Z, Hilder L & Sullivan E 2013. Australia's mothers and babies 2011. Cat. no. PER 59. Canberra: AIHW.

COAG (Council of Australian Governments) 2008. Intergovernmental Agreement on Federal Financial Relations: National Indigenous Reform Agreement. Canberra: Commonwealth of Australia. Viewed 5 May 2019. http://www.federalfinancialrelations.gov.au/content/npa/health/_archive/indigenous-reform/national-agreement_sept_12.pdf.

COAG 2018. Communique: Meeting of the Council of Australian Governments, Adelaide 12 December 2018.

COAG 2018. Communique: Meeting of the Council of Australian Governments. Adelaide 12 December 2018. Viewed 5 May 2019. <https://www.coag.gov.au/sites/default/files/communique/coag-communique-dec-2018.pdf>.

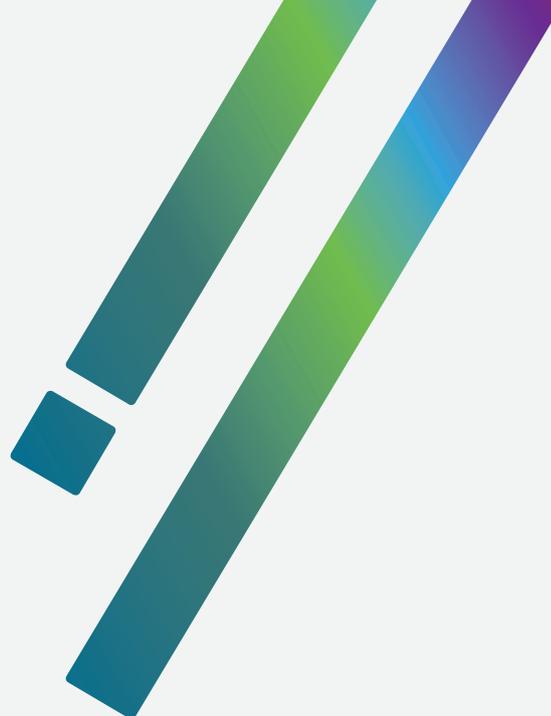
PM&C (Prime Minister & Cabinet) 2015. Closing the Gap: Prime Minister's report 2015. Viewed 5 May 2019. https://www.pmc.gov.au/sites/default/files/publications/Closing_the_Gap_2015_Report_1.pdf Canberra: PM&C.

List of tables

Table 5.1: Reclassified Indigenous deaths, by sex and age, 2001–2015 (%).....	9
Table 5.2: Indigenous life expectancy at birth, adjusted and not adjusted for misclassification, by sex, 2001–2005, 2006–2010, 2011–2015.....	10
Table 6.1: Indigenous life expectancy at birth based on 2 denominator populations, by sex, 2001–2005, 2006–2010, 2011–2015.....	12
Table 6.2: Estimates of Indigenous population based on the Census, by sex, at 30 June 2001, 2006, 2011, 2016	12
Table A1.1: Key studies towards improving Indigenous identification in data.....	14

List of figures

Figure 4.1: Misclassified Indigenous deaths as a percentage of records originally identified as Indigenous on the NDI, by sex and age group, 2001–2005, 2006–2010, 2011–2015.....	7
Figure 5.1: Reported and enhanced Indigenous deaths, by age and sex, Australia, 2001–2015	8
Figure 6.1: Life expectancy at birth, by age, sex, and Indigenous status, 2001–2005, 2006–2010, 2011–2015.....	10



Measuring progress on Aboriginal and Torres Strait Islander health and welfare outcomes relies on consistent, complete, and reliable identification of Indigenous Australians in key data collections.

Incomplete and inconsistent reporting of Indigenous identification occurs through a combination of Indigenous misclassification by data providers at the point of data collection, and Indigenous people choosing not to identify as Indigenous in certain circumstances.

This report provides an overview of AIHW work on improving Indigenous identification, particularly in enhancing mortality estimation through statistical data linkage.

aihw.gov.au



Stronger evidence,
better decisions,
improved health and welfare

