Using linkage between hospital and ABS mortality data to enhance reporting of deaths among Aboriginal and Torres Strait Islander peoples

Sarah E. Neville, Lee K. Taylor, Helen Moore  
Centre for Epidemiology and Research, NSW Department of Health

Richard Madden  
National Centre for Classification in Health, University of Sydney

Ian Ring  
Centre for Health Service Development, University of Wollongong, New South Wales

Lisa Jackson Pulver  
School of Public Health and Community Medicine, University of New South Wales

Leonie Tickle  
Department of Actuarial Studies, Macquarie University, New South Wales

Austrailian Bureau of Statistics (ABS) mortality data is the main source of information about the mortality of Aboriginal and Torres Strait Islander peoples in Australia. Reporting of deaths among Aboriginal and Torres Strait Islander peoples on these data is incomplete. Using indirect methods, ABS has estimated the level of enumeration in NSW from 2002 to 2006 to be 45%, based on comparison between deaths registered among Aboriginal and Torres Strait Islander peoples and expected numbers of deaths from ABS population projections.1

More recently, the ABS Indigenous Mortality Quality Study used record linkage between 2006 Census records and deaths registered during the 11-month period following the Census to estimate the level of enumeration of deaths among Aboriginal and Torres Strait Islander peoples in NSW at 76.3%.2

Record linkage provides a possible mechanism for enhancing the reporting of Aboriginal and Torres Strait Islander peoples on routinely collected data, including mortality data, by using information from all linked datasets to update information on whether a person is Aboriginal or Torres Strait Islander on any one of the linked datasets. This would allow the production of better quality ‘enhanced’ health status measures, including mortality measures. As a result, routinely collected data will be better able to inform policy, health services planning, health promotion and evaluation of these activities.

Data linkage has been used to improve the reporting of Aboriginal and Torres Strait Islander peoples on routinely collected mortality data in Australia, New Zealand and Canada.3 In Australia, in addition to the Indigenous Mortality Quality Study, a recently published Western Australian study used data linkage to improve the quality of mortality estimates regarding the Aboriginal and Torres Strait Islander population. That study’s major aim was to use data linkage to reduce the number of records where the

Abstract

Objective: To investigate the potential of record linkage between the Australian Bureau of Statistics (ABS) mortality data and the NSW Admitted Patient Data Collection (APDC) to improve reporting of deaths among Aboriginal and Torres Strait Islander peoples.

Methods: ABS mortality data for 2002 to 2006 were linked with APDC records for 2001 to 2006. Six algorithms were developed to enumerate deaths. Possible biases by age, sex and geographic remoteness were investigated.

Results: Levels of reporting ranged from baseline reporting on the ABS mortality data to the largest enhancement with the ‘ever reported as Aboriginal or Torres Strait Islander’ algorithm. Enhancement was more likely in females, older people and residents of major cities.

Conclusions: Data linkage substantially improved reporting of Aboriginal and Torres Strait Islander deaths. An algorithm that includes both the number of APDC records and the number of facilities reporting a person as Aboriginal or Torres Strait Islander was considered most promising.

Implications: Inclusion of other datasets in the enhancement process is warranted to further improve reporting and address possible bias produced by using APDC records only. Further work should take into account the possibility that a person may be falsely reported as Aboriginal or Torres Strait Islander or not reported in either hospital or death records.

Key words: data linkage, Aboriginal, Indigenous, mortality.


Submitted: August 2010  
Revision requested: December 2010  
Accepted: March 2011

Correspondence to:  
Dr Lee K Taylor, Centre for Epidemiology and Research, NSW Department of Health.  
Locked Bag 961, North Sydney, NSW; e-mail: ltayl@doh.health.nsw.gov.au
relevant information was missing. Death records were linked with the Western Australian Hospital Morbidity Database System, Mental Health Information System and Midwives Notification System. It was found that the increasing proportion of deaths where the unlinked records did not indicate whether a person was Aboriginal or Torres Strait Islander, over time, led to an overestimate of improvements in life expectancy for Aboriginal and Torres Strait Islander peoples.⁴

In New Zealand an ‘ever Māori’ method of ethnicity classification has been developed using record linkage to improve underenumeration of Māori persons in health datasets, including mortality data.⁵ In Canada, Indigenous status is not usually captured on administrative health datasets or on mortality data. Instead, routinely collected data is linked with the Indian Register in order to produce estimates for the population covered by the Register.³

This study aims to estimate the level of improvement in reporting of deaths among Aboriginal and Torres Strait Islander peoples in NSW on the ABS mortality data using record linkage with the NSW Admitted Patient Data Collection (APDC) and to investigate possible sources of bias caused by the record linkage.

Methods

Data sources

The APDC covers demographic and episode-related data for every inpatient that is separated from any public, private or and repatriation hospital, private day procedure centre or public nursing home in NSW.

ABS mortality data is based on year of registration of death, and includes both coronial and non-coronal deaths. Deaths are reported to the Registry of Births, Deaths and Marriages (RBDM), which forwards the details to the ABS, where the information is checked and coded. NSW death records available for linkage include deaths that occurred in NSW but not deaths of NSW residents who died interstate.

The APDC was selected for linkage with the ABS mortality data because it is a large dataset covering all ages, patients must be asked whether they are of Aboriginal or Torres Strait Islander origin at each hospital admission,⁶ and reporting is more accurate than for the mortality data – it has been estimated that 88% of NSW APDC records that relate to Aboriginal or Torres Strait Islander persons were correctly reported in 2007.⁷

Linked APDC and ABS mortality data were drawn from the Centre for Health Record Linkage (CHeReL) Master Linkage Key, using the best practice approach in privacy preserving record linkage.⁸ The CHeReL used the following information to link APDC records for the same person: full name; address; sex; date of birth; country of birth; hospital code; medical record number; dates of admission and discharge; hospital transferred to; hospital transferred from; and date of death (if the patient died in hospital). RBDM death registration records were linked to APDC records using full name, address, sex, date of birth and date of death. Linkage was carried out using the open source probabilistic record linkage software ChoiceMaker.⁹ ABS mortality data were deterministically linked to RBDM death registration records using the death registration number. Information about Indigenous status was not used for record linkage. Quality assurance procedures at the CHeReL ensured that there were no more than 5/1,000 missed links and no more than 5/1,000 false positive links in the CHeReL Master Linkage Key.¹⁰

The final dataset was made up of ABS mortality data for all deaths registered in NSW in the period 2002 to 2006 (n=227,205 records), and linked APDC records for the period 2001 to 2006 (n=12,809,691 records).

The NSW Population and Health Services Research Ethics Committee approved the study.

Analysis

Both the ABS mortality data and the APDC report people as belonging to one of the following groups: Aboriginal; Torres Strait Islander; both Aboriginal and Torres Strait Islander; neither Aboriginal nor Torres Strait Islander; or not stated. For the purpose of this study, the first three groups were condensed into one group for Aboriginal or Torres Strait Islander peoples.

Initially, the consistency of reporting of hospital admissions for Aboriginal and Torres Strait Islander peoples on the APDC was explored both within and across facilities for all people with at least one record where the person was reported as Aboriginal or Torres Strait Islander. Within facility consistency was defined as agreement on all records for a person within a facility. Complete consistency was defined as agreement on all records for a person both within and across facilities.

To address the false negative reporting of Aboriginal and Torres Strait Islander peoples in the mortality data, we enhanced reporting by defining algorithms to extract the relevant information for persons represented on the linked mortality and APDC data. These algorithms are:

• Algorithm 1 used reporting of deaths among Aboriginal and Torres Strait Islander peoples as defined by the ABS mortality data. This provides a baseline for comparison for the remaining five algorithms.
• Algorithm 2 uses an ‘ever reported as Aboriginal or Torres Strait Islander’ criterion. If an individual was reported as Aboriginal or Torres Strait Islander on any linked record (i.e. mortality or APDC records) then the person was designated as Aboriginal or Torres Strait Islander. This algorithm represents the upper bound or maximum level of enhancement achievable from linkage of mortality and APDC data.
• Algorithm 3 uses a proportional record-level criterion for APDC records. If an individual was reported as Aboriginal or Torres Strait Islander on the mortality record or on at least 50% of their APDC records then the person was designated as Aboriginal or Torres Strait Islander.
• Algorithm 4 uses a proportional facility-level criterion for APDC records. If an individual was reported as Aboriginal or Torres Strait Islander on the mortality record or on at least 50% of their most recent APDC records from each facility visited then the person was designated as Aboriginal or Torres Strait Islander.
• Algorithm 5 uses proportional criteria at both the record and facility levels for APDC records. If an individual was reported as Aboriginal or Torres Strait Islander on the mortality record or on 50% or more of records in 50% or more of facilities recorded in the APDC data then the person was designated as Aboriginal or Torres Strait Islander.

• Algorithm 6 departs from a proportional methodology, and is a rule defined by a specified number of records. Again we have a two-stage algorithm, defined at both within and cross-facility levels. If an individual was either: reported as Aboriginal or Torres Strait Islander on the mortality record, or in two or more records in two or more facilities on the APDC, then the person was designated as Aboriginal or Torres Strait Islander. If only one or two records in one or two facilities were available for the person, one record in one facility was considered sufficient. Each algorithm was applied in two ways:
  1. Each year of mortality data was enhanced using both the current and previous calendar year of APDC data. For example, the 2006 mortality data was enhanced with all APDC data where the year of hospital separation was either 2005 or 2006.
  2. Mortality data for 2006 was enhanced using 2, 3, 4, 5 and 6 years of APDC data.

The possibility of false positive reports was explored by ascertaining the number of persons who were reported as Aboriginal or Torres Strait Islander on the mortality data, but were consistently reported as neither Aboriginal nor Torres Strait Islander on their linked APDC records.

Numbers of deaths for 2004 to 2006 resulting from algorithms 1, 2 and 6 were further examined by age, sex and geographic remoteness to explore possible sources of bias in the enhancement process. Percentage enhancement was calculated for algorithms 2 and 6 by expressing the difference between the number of enhanced deaths and the baseline as a fraction of the baseline number of deaths. The enhanced Accessibility/Remoteness Index of Australia (ARIA+) was used to measure geographic remoteness.

The ‘ever reported as Aboriginal or Torres Strait Islander’ definition of algorithm 2 provides the highest level of enhancement (29.7% for 2006 data) and the highest number of deaths. Algorithm 6, which relies on a specified number of records, produces the second largest enhancement (23.6% for 2006 data). Algorithms 3, 4 and 5, which are based on proportions of records reported as relating to Aboriginal or Torres Strait Islander persons, produced intermediate results (22.6%, 22.8%, and 21.7% respectively for 2006 data).

The impact of using varying numbers of years of APDC data to enhance was assessed by enhancing the 2006 mortality data with between 2 and 6 years of APDC records. The algorithms based on proportions (3, 4 and 5) produced fairly consistent enhancement (23.6%, 22.8%, and 21.7% respectively for 2006 data).

Table 1: Number of deaths among Aboriginal and Torres Strait Islander peoples in ABS mortality data by enhancement algorithm, NSW, 2002 to 2006.

| Year | Algorithm 1 Any ABS mortality record | Algorithm 2 Any ABS mortality record or linked APDC record | Algorithm 3 ABS mortality record or at least 50% of linked APDC records | Algorithm 4 ABS mortality record or at least 50% most recent linked APDC records from each facility | Algorithm 5 ABS mortality record or at least 50% linked APDC records in at least 50% facilities | Algorithm 6 ABS mortality record or 2 or more linked APDC records in 2 or more facilities
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>507</td>
<td>636</td>
<td>579</td>
<td>594</td>
<td>584</td>
<td>598</td>
</tr>
<tr>
<td>2003</td>
<td>479</td>
<td>619</td>
<td>563</td>
<td>586</td>
<td>581</td>
<td>587</td>
</tr>
<tr>
<td>2004</td>
<td>470</td>
<td>612</td>
<td>560</td>
<td>572</td>
<td>569</td>
<td>580</td>
</tr>
<tr>
<td>2005</td>
<td>497</td>
<td>638</td>
<td>584</td>
<td>594</td>
<td>588</td>
<td>602</td>
</tr>
<tr>
<td>2006</td>
<td>508</td>
<td>659</td>
<td>623</td>
<td>624</td>
<td>618</td>
<td>628</td>
</tr>
</tbody>
</table>

Source: Linked ABS mortality data and NSW Admitted Patient Data Collection (HOIST). Centre for Epidemiology and Research, NSW Department of Health.

Notes:
a. Two years of admitted patient data used to enhance each year of ABS mortality data.
b. Algorithm 6: where only one or two records in one or two facilities available, one record or facility is considered sufficient.

Results

Of all APDC records in the linked dataset, 97.3% of records were reported as relating to non-Aboriginal or Torres Strait Islander persons, 2.0% were recorded for Aboriginal or Torres Strait Islander persons and 0.7% were recorded as ‘not stated’. For the mortality data, 97.3% of records were reported as relating to non-Aboriginal or Torres Strait Islander persons, 1.1% were recorded for Aboriginal or Torres Strait Islander persons and 1.6% were recorded as ‘not stated’.

Within facility consistency (for APDC records from 2005 and 2006) for persons with at least one APDC record where the person was reported as Aboriginal or Torres Strait Islander was 86.6%. In contrast, complete consistency was 64.3%.

The results of the six algorithms using two years of APDC data are shown in Table 1. Algorithm 1 creates a baseline using the un-enhanced mortality data and produces the lowest number of deaths. The ‘ever reported as Aboriginal or Torres Strait Islander’ definition of algorithm 2 provides the highest level of enhancement (29.7% for 2006 data) and the highest number of deaths. Algorithm 6, which relies on a specified number of records, produces the second largest enhancement (23.6% for 2006 data). Algorithms 3, 4 and 5, which are based on proportions of records reported as relating to Aboriginal or Torres Strait Islander persons, produced intermediate results (22.6%, 22.8%, and 21.7% respectively for 2006 data).

The impact of using varying numbers of years of APDC data to enhance was assessed by enhancing the 2006 mortality data with between 2 and 6 years of APDC records. The algorithms based on proportions (3, 4 and 5) produced fairly consistent enhancement (23.6%, 22.8%, and 21.7% respectively for 2006 data).
enhancements regardless of the number of years of APDC data used for enhancement. For example, the number of deaths reported among Aboriginal and Torres Strait Islander peoples resulting from algorithm 3 varied between 618 and 625 using different numbers of years of data to enhance reporting. In contrast, numbers of deaths of Aboriginal and Torres Strait Islander peoples created by algorithms 2 and 6 continued to increase with the number of years of linked APDC data. Deaths identified using the ‘ever reported as Aboriginal or Torres Strait Islander’ definition of algorithm 2 increased from 659 to 738 using 2 and 6 years of APDC records respectively. There was a less marked increase in the number of deaths resulting from algorithm 6, rising from 628 to 642 using 2 and 6 years of APDC records, respectively.

Of the 508 deaths of Aboriginal and Torres Strait Islander peoples reported to ABS in 2006, 89 (17.5%) did not link with the current or previous year of APDC data. Of the remaining 419 deaths that did link to at least one APDC record, 129 deaths were linked to APDC records where the person was consistently reported as non-Aboriginal or Torres Strait Islander. The majority of these 129 deaths (80%) had five or fewer linked APDC records, and more than half (53%) had visited only one facility. One individual stood out with greater than 10 APDC records reported from greater than 10 facilities.

Eighteen deaths of Torres Strait Islander people were reported in the 2006 mortality data, and an additional five deaths were identified by the linked APDC records from 2005 and 2006 using the ‘ever reported as Aboriginal’ algorithm, bringing the enhanced number of reported deaths among Torres Strait Islander people to 23.

Demographic breakdowns were produced to assess the potential bias introduced by enhancing mortality data with APDC records alone. Table 2 presents the number of deaths reported among Aboriginal or Torres peoples for 2004 to 2006 by 5 year age group for algorithms 1, 2 and 6. The level of enhancement varied between age groups, ranging from 12% in 35 to 39 year olds to 63% in children between 5 and 9 years for the ‘ever reported as Aboriginal or Torres Strait Islander’ algorithm 2.

Enhancement also differed between males and females. Females tended to have higher levels of enhancement, with 31.8% and 24.8% enhancement for algorithms 2 and 6 respectively. In contrast, deaths in males were enhanced by 27.6% using algorithm 2, and 21.1% using algorithm 6.

The level of enhancement varied by geographic area (Table 3). Percentage enhancement for both algorithms 2 and 6 was highest in major cities, and generally decreased as areas become more remote. The fall in the level of enhanced reporting was most marked when moving from outer regional to remote areas, where the percentage enhancement for algorithms 2 and 6 decreased by 14 and 11 percentage points respectively.

Table 2: Number of deaths among Aboriginal and Torres Strait Islander peoples in ABS mortality data by age and enhancement algorithm, NSW, 2004 to 2006.a

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Algorithm 1 Baseline ABS mortality record</th>
<th>Algorithm 2 Any ABS mortality or linked APDC record</th>
<th>Algorithm 6 Any ABS mortality record or 2 or more linked APDC records in 2 or more facilitiesb</th>
<th>% enhancement for algorithm 2</th>
<th>% enhancement for algorithm 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>81</td>
<td>104</td>
<td>103</td>
<td>28.4</td>
<td>27.2</td>
</tr>
<tr>
<td>5-9</td>
<td>8</td>
<td>13</td>
<td>11</td>
<td>62.5</td>
<td>37.5</td>
</tr>
<tr>
<td>10-14</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>20.0</td>
<td>20.0</td>
</tr>
<tr>
<td>15-19</td>
<td>21</td>
<td>25</td>
<td>23</td>
<td>19.0</td>
<td>9.5</td>
</tr>
<tr>
<td>20-24</td>
<td>27</td>
<td>34</td>
<td>34</td>
<td>25.9</td>
<td>25.9</td>
</tr>
<tr>
<td>25-29</td>
<td>31</td>
<td>40</td>
<td>37</td>
<td>29.0</td>
<td>19.4</td>
</tr>
<tr>
<td>30-34</td>
<td>57</td>
<td>67</td>
<td>66</td>
<td>17.5</td>
<td>15.8</td>
</tr>
<tr>
<td>35-39</td>
<td>83</td>
<td>93</td>
<td>93</td>
<td>12.0</td>
<td>12.0</td>
</tr>
<tr>
<td>40-44</td>
<td>94</td>
<td>114</td>
<td>109</td>
<td>21.3</td>
<td>16.0</td>
</tr>
<tr>
<td>45-49</td>
<td>92</td>
<td>118</td>
<td>116</td>
<td>28.3</td>
<td>26.1</td>
</tr>
<tr>
<td>50-54</td>
<td>117</td>
<td>140</td>
<td>135</td>
<td>19.7</td>
<td>15.4</td>
</tr>
<tr>
<td>55-59</td>
<td>99</td>
<td>123</td>
<td>118</td>
<td>24.2</td>
<td>19.2</td>
</tr>
<tr>
<td>60-64</td>
<td>162</td>
<td>212</td>
<td>202</td>
<td>30.9</td>
<td>24.7</td>
</tr>
<tr>
<td>65-69</td>
<td>133</td>
<td>174</td>
<td>162</td>
<td>30.8</td>
<td>21.8</td>
</tr>
<tr>
<td>70-74</td>
<td>141</td>
<td>187</td>
<td>176</td>
<td>32.6</td>
<td>24.8</td>
</tr>
<tr>
<td>75-79</td>
<td>111</td>
<td>150</td>
<td>135</td>
<td>35.1</td>
<td>21.6</td>
</tr>
<tr>
<td>80-84</td>
<td>88</td>
<td>123</td>
<td>116</td>
<td>39.8</td>
<td>31.8</td>
</tr>
<tr>
<td>85-89</td>
<td>67</td>
<td>96</td>
<td>87</td>
<td>43.3</td>
<td>29.9</td>
</tr>
<tr>
<td>90+</td>
<td>58</td>
<td>90</td>
<td>81</td>
<td>55.2</td>
<td>39.7</td>
</tr>
<tr>
<td>Total</td>
<td>1,475</td>
<td>1,909</td>
<td>1,810</td>
<td>29.4</td>
<td>22.7</td>
</tr>
</tbody>
</table>

Source: Linked ABS mortality data and NSW Admitted Patient Data Collection (HOIST). Centre for Epidemiology and Research, NSW Department of Health.
Notes: a. ABS records for 2004, 2005 and 2006 each enhanced with two years of admitted patient data.
b. Algorithm 6: where only one or two records in one or two facilities available, one record or facility is considered sufficient.
Discussion

We used six algorithms to estimate the number of deaths among Aboriginal and Torres Strait Islander persons, representing a spectrum of reporting from the baseline of using the un-enhanced mortality data, through to the ‘ever reported as Aboriginal or Torres Strait Islander’ criterion of algorithm 2. The level of enhancement reflected whether the algorithm depended on a specified number of records reported as belonging to an Aboriginal or Torres Strait Islander person, or whether the algorithm was based on a proportion of records. We also explored the extent to which the level of enhancement varied with the number of years of APDC data used in the enhancement process.

Algorithms 2 and 6 result in the highest level of enhanced reporting. However, the ‘ever reported as Aboriginal or Torres Strait Islander’ criterion is difficult to justify conceptually. In the extreme case, a person with a death record indicating that they are non-Aboriginal or Torres Strait Islander and numerous linked APDC records but only one APDC record indicating they are Aboriginal or Torres Strait Islander will be regarded as Aboriginal or Torres Strait Islander for the purposes of enhanced reporting. Algorithm 6 provides some resolution for this problem.

Algorithm 6 requires two or more records within a facility (or at least one record if there are only one or two records for the facility) to be reported as relating to an Aboriginal or Torres Strait Islander person in order for the person to be regarded as Aboriginal or Torres Strait Islander within that facility. Furthermore, if an individual has visited two or more facilities, they must be reported as Aboriginal or Torres Strait Islander in at least two of these facilities in order to be considered Aboriginal or Torres Strait Islander overall. This algorithm is more robust to extreme scenarios than the ‘ever reported as Aboriginal or Torres Strait Islander’ algorithm 2. The three algorithms, 3, 4 and 5, that use proportions as their central methodology result in fairly similar levels of enhancement for the five years of mortality data. We consider algorithms 5 and 6 as the most methodologically sound for use in enhanced reporting of mortality data using APDC records. Making a choice between these two will depend on what is considered a ‘good’ attribute of an enhancement algorithm. One could argue that under-reporting of Aboriginal and or Torres Strait Islander peoples on the APDC is a greater problem than the risk of false-positive reporting and, therefore, a specified number of records reported, such as n=2 in algorithm 6, rather than a majority, is reasonable.

We found that there was little difference in the resulting number of deaths when we used two or six years of APDC data for enhancement using the proportional algorithms (3, 4 and 5). For algorithm 6, which uses a rule defined by a specified number of records, there was a modest increase in the resulting numbers of deaths, from 628 to 642 using two and six years of APDC records, respectively. If linked APDC and mortality data are to be used for monitoring trends in deaths among Aboriginal and Torres Strait Islander peoples overall, the same number of years of APDC data should be used to enhance each year of mortality data. Any decision about the optimum number of years of APDC data used for enhancement will, therefore, involve a balance between correcting the under-enumeration of deaths among Aboriginal or Torres Strait Islander for a given year versus the number of years of trend information desired and the number of years of linked data available. For trend analyses, our results suggest that two years of APDC data would probably be sufficient.

We found that, although the highest percentage enhancement occurred in children aged between five and nine, this was followed closely by those people aged 85 years and over. This pattern may be caused by older people being more likely to be hospitalised. Furthermore, we found that enhancement was more likely to occur in females than males, possibly related to females being more likely to have a hospital record due to pregnancy and childbirth. It, therefore, seems likely that enhanced reporting of Aboriginal or Torres Strait Islander peoples on mortality data by linkage with the APDC data may result in differential improvement in reporting among females and the elderly, which should be borne in mind as a source of possible bias.

Table 3: Number of deaths among Aboriginal and Torres Strait Islander peoples in ABS mortality data by ARIA+ category and enhancement algorithm, NSW, 2004 to 2006.a

<table>
<thead>
<tr>
<th>ARIA category</th>
<th>Algorithm 1 Baseline ABS mortality record</th>
<th>Algorithm 2 Any ABS mortality or linked APDC record</th>
<th>Algorithm 6 ABS mortality record or 2 or more linked APDC records in 2 or more facilities</th>
<th>% enhancement for algorithm 2</th>
<th>% enhancement for algorithm 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Cities</td>
<td>491</td>
<td>710</td>
<td>658</td>
<td>44.5</td>
<td>33.9</td>
</tr>
<tr>
<td>Inner Regional</td>
<td>469</td>
<td>582</td>
<td>555</td>
<td>24.2</td>
<td>18.3</td>
</tr>
<tr>
<td>Outer Regional</td>
<td>332</td>
<td>411</td>
<td>394</td>
<td>23.9</td>
<td>18.9</td>
</tr>
<tr>
<td>Remote</td>
<td>120</td>
<td>132</td>
<td>130</td>
<td>9.6</td>
<td>7.7</td>
</tr>
<tr>
<td>Very Remote</td>
<td>33</td>
<td>34</td>
<td>34</td>
<td>4.5</td>
<td>4.2</td>
</tr>
<tr>
<td>Total</td>
<td>1,475</td>
<td>1,909</td>
<td>1,810</td>
<td>29.4</td>
<td>22.7</td>
</tr>
</tbody>
</table>

Source: Linked ABS mortality data and NSW Admitted Patient Data Collection (HOIST). Centre for Epidemiology and Research, NSW Department of Health.

Notes:  
a. ABS records for 2004, 2005 and 2006 each enhanced with 2 years of admitted patient data.  
b. Total include records with interstate or missing SLAs.  
c. Algorithm 6: where only one or two records in one or two facilities available, one record or facility is considered sufficient.
The substantial differences in enhanced reporting resulting from
the breakdown by ARIA+ illustrate the potential of record linkage
to improve the quality of reporting of Aboriginality in major cities
and inner and outer regional areas.

This study has extended previous work carried out in both
Western Australia and New Zealand. Draper et al. used data linkage
to improve mortality estimates by replacing missing information
about Aboriginal and Torres Strait Islander persons with information
yielded by the linkage process. This process does not take into
account the situation where persons are falsely reported as non-
Aboriginal or Torres Strait Islander. Our study only allowed records
that were reported as relating to ‘non-Aboriginal or Torres Strait
Islander’ persons to be updated by the enhancement process.

This study goes beyond the ‘ever Māori’ method of ethnicity
classification developed by the New Zealand Ministry of Health.
The algorithms developed include rules that are more selective
and probably more robust than their ‘ever reported as Aboriginal
or Torres Strait Islander’ counterpart.

There are some other issues to consider in designing algorithms
to utilise linked data that we have not incorporated into the six
algorithms described in this study. One of these issues is that of
potential false positive reporting. All the algorithms examined in
this paper assumed that reporting of Aboriginal and Torres Strait
Islander peoples on the mortality data was correct. We found a
small number of persons who were reported as Aboriginal or Torres
Strait Islander on the mortality data but had multiple records on the
APDC where the person was reported as non-Aboriginal or Torres
Strait Islander. It is not possible to know how many of these are
false positive reports in the first instance or incorrect links due to
the probabilistic approach used for record linkage and how many
are correctly reported in the mortality set but not in the APDC.
The quality assurance procedures at the CHeReL allow a maximum of
5/1,000 links to be false positives; so it is possible that there are
some false positive links in the data used for this study. We were
unable to identify any studies that have examined false-positive
reporting of Aboriginal and Torres Strait Islander peoples on the
ABS mortality data.

The enhancement algorithms we used would not include
Aboriginal or Torres Strait Islander persons who were not reported
as such on their death records and who were not admitted to hospital,
or who were admitted to hospital but never recorded as Aboriginal
or Torres Strait Islander. In order to reduce biases introduced by
using APDC records for enhancement, other administrative health
datasets could be incorporated into the linkage and enhancement
process. This would lead to more complex algorithms that were not
explored in this study.

We have described the potential of five different algorithms
to enhance the reporting of Aboriginality in mortality data, using linked
APDC records. We suggest an algorithm that takes into account the
number of records and the number of facilities that report a person
as Aboriginal or Torres Strait Islander on the APDC records is likely
to give the most reliable results.

Ultimately, accurate information about Aboriginal and Torres
Strait Islander peoples on health data depends on the correct
information being collected in the first instance. The number of false
negative reports would be reduced if patients were always asked a
question about whether they are Aboriginal or Torres Strait Islander
at the time health information is being collected. Record linkage is
not a substitute for this.

However, record linkage does have the potential to substantially
improve health measures for Aboriginal and Torres Strait Islander
peoples that are based on mortality data that have already been
collected, such as life expectancy, and also for measures of
healthcare utilisation that are derived from other routinely collected
health data, particularly those relating to hospitalisations and
emergency department attendances. Further work is needed, first
to explore the use of multiple sources of administrative data in the
linkage and enhancement process to reduce false negative reports
of Aboriginality, second to deal with the possibility of false positive
reports of Aboriginality on the mortality data, and, finally, to improve
identification in both mortality and APDC records.

Acknowledgements

The authors thank the AHMRC for its support of the project,
and acknowledge the HOIST team in the NSW Department of
Health for hosting the linked datasets. ChoiceMaker Technologies
Inc. developed the ChoiceMaker software and contributed it to the
open source community.

References

2007.
2. Australian Bureau of Statistics. 4723.0 – Information Paper: Census Data
2008.
Indigenous status on mortality estimates? An assessment using record linkage
6. NSW Department of Health. Aboriginal and Torres Strait Islander Origin –
Recording of Information of Patients and Clients. Policy Directive Number
2005_547 [Internet]. Sydney (AUST): State Government of New South Wales;
7. Australian Institute of Health and Welfare. Indigenous Identification in Hospital
Separations Data – Quality Report. Health Services Series No.: 35. Canberra
[Internet]. Brooklyn (NY): ChoiceMaker Technology, 2011 [cited 2011 Jan
10. Centre for Health Record Linkage. Quality Assurance in Record
Linkage [Internet]. Alexandria (AUST): Cancer Institute NSW, CHeReL;
CHeReL.QualityAssuranceJuly2008.pdf
11. GISCA. About ARIA+ – Accessibility/Remoteness Index of Australia [Internet].