

XXVI IUSSP International Population Conference, Marrakech, 27 Sept. to 2 Oct. 2009

Session 198 - Improving the usefulness of 2010 Population Census round (International Statistical Institute)

USING POPULATION CENSUS AND DEATH REGISTRATION DATA LINKAGE IN COMPILING LIFE EXPECTANCY ESTIMATES FOR INDIGENOUS AUSTRALIANS

Patrick M. Corr, Australian Bureau of Statistics

Abstract

Life expectancy estimates for Aboriginal and Torres Strait Islander (Indigenous) Australians are important in assessing Indigenous disadvantage. The compilation of accurate life tables for Indigenous Australians presents particular difficulties due to data quality limitations in both Indigenous population estimates and death registrations. The application of standard life table methods without any adjustments can lead to results with numerator/denominator bias.

The Australian Bureau of Statistics used data linkage techniques with the 2006 Australian Population Census to derive Indigenous identification consistency ratios for Indigenous death registrations compiled from vital event registers. This paper presents a summary of the life expectancy results from direct adjustment of the number of Indigenous-identified registered deaths for differential Indigenous identification using the results of the data linkage study and Census-based population estimates adjusted for net undercount and misclassification error. Comparative results are presented from selected other methods and approaches.

Introduction

Addressing social disadvantage in respect of Aboriginal and Torres Strait Islander Australians has been, and continues to be, a key policy issue for all levels of government in Australia. Since 2003 an *Overcoming Indigenous Disadvantage* report, developed at the request of the Council of Australian Governments, has been prepared on a biennial basis. The report provides indicators of Indigenous disadvantage and where possible it measures how that disadvantage has changed over time.

The Council of Australian Governments has identified a number of targets related to Indigenous well-being. These are to:

- a) Close the life expectancy gap between Indigenous and non-Indigenous Australians within a generation
- b) Halve the gap in mortality rates for Indigenous children under five within a decade
- c) Halve the gap for Indigenous students in reading, writing and numeracy within a decade
- d) Halve the gap in employment outcomes for Indigenous Australians within a decade
- e) Provide access for all four year olds in remote communities to early childhood education within five years
- f) At least halve the gap for Indigenous students in Year 12 attainment or equivalent attainment rates by 2020.

The focus on Indigenous disadvantage and the desire to quantify the extent, and to measure the change in the level, of disadvantage has brought the quality of data about Indigenous Australians, in particular demographic statistics such as life expectancy at birth, and the underlying data sources under close scrutiny.

A by-product of the strong demand for data is that Indigenous Australians have a very high reporting workload in terms of providing information for the purpose of government administration, surveys or independent research. Furthermore, the variability in reporting of 'Indigenous status' in administrative and survey (including population census) documents has provided a strong impetus for linking these

different data sources so as to further enhance the value of the data to better inform on the extent of disadvantage.

The balance of this paper examines some of the primary data sources for compiling Indigenous life expectancy data and the data quality issues associated with them. The paper discusses the use of data linking across administrative and statistical collection sources to improve the quality of Indigenous life expectancy estimates, and explores some of the steps taken by ABS prior to implementing its Census Data Enhancement Project so as to address privacy concerns associated with data linking.

Some context

Australia, like most countries has a wide range of administrative data. Births, deaths and marriages are required to be registered with the relevant state or territory Registrars of Births, Deaths and Marriages; persons are required to register for access to the national health scheme (Medicare), persons seeking social security benefits such as unemployment benefits, aged pension, and other family benefits, etc must be registered with the relevant government agency (Centrelink), persons need to register or enrol with the Australian Electoral Commission to vote in national elections; etc. However, unlike some other countries, Australia does not maintain a population register and with the exception of fraud detection purposes, there is very limited linking of these registers. Over time there has been vocal opposition from the Australian public to the establishment of population registers and associated identification cards, and at the same time a strengthening in the privacy rights of individuals.

Furthermore, in respect of Aboriginal and Torres Strait islander Australians – the Indigenous peoples of Australia - there is often a question on the necessary registration forms as to the Indigenous status of persons, however, the reporting of Indigenous status is quite variable.

It is within this context the paper is set.

Data sources

Life expectancy estimates are calculated by constructing a life table based on age-sex specific death rates. These rates require quality data on both the number of deaths by age and sex for the population being analysed, and the population at risk of dying for each age group and sex. When compiling life tables for sub-populations, an additional dimension is introduced being the identification of the sub-population group in the death estimates, population estimates, and the birth estimates (which are used as the denominator for infant mortality rates).

The key data sources for the compilation of a life table for Indigenous Australians are:

- a) the 5 yearly Census of Population and Housing conducted by the ABS;
- b) the Post Enumeration Survey conducted by ABS immediately after the Census;
- c) Death registrations sourced from the Registrars of Births, Deaths and Marriages; and
- d) Birth registrations sourced from the Registrars of Births, Deaths and Marriages

The first sources provide the population benchmarks at 5 yearly intervals. The latter two elements provide estimates of births and deaths. The other component to Indigenous population change (ie migration) is assumed to be minimal in the Australian situation.

For each of the above data sources there is an Indigenous status question to the effect of:

“Is the person of Aboriginal or Torres Strait Islander origin”

Response categories allow for Aboriginal, Torres Strait Islander, both Aboriginal and Torres Strait Islander, or no.

While there are defining characteristics to be classified as an Indigenous Australian, the reality is that for the above data sources, the results are dependent on the perception of the person completing the form, and it is often the case that the person completing the form is not the person about whom the information is being sought. The other reality is that the question is not always answered.

Data quality

In relation to each of the above data sources there are a range of quality issues.

The quality of the 5 yearly benchmarks from the Census of Population and Housing are compromised by the variable coverage of Indigenous Australians. The level of net undercoverage of Indigenous Australians (11.5% in 2006) is markedly higher than for the total Australian population (2.7% in 2006). In turn the estimates of net undercount which are sourced from the Census Post Enumeration Survey (PES) conducted directly after the Census are subject to sampling error. The net undercount estimate of 11.5% has a standard error of 2.3 percentage points, that is, the 95% confidence interval is a range between 6.9% and 16.1%

Sampling error is not the only aspect that impacts on the quality of the census benchmark data. Non sample error also presents in a number of forms. The reporting of Indigenous status is quite variable over time. This reflects both a changing propensity of people to identify as Indigenous Australians and also misreporting of Indigenous status. The response of individuals to questions as to Indigenous status can be influenced by a range of factors, including the fact that the question is often completed by another family or household member. As well, individuals' perceptions including personal and cultural aspects associated with identifying as Indigenous, will influence how they might respond.

To illustrate this, it is worth examining the different responses to the Indigenous status question in the 2006 Census of Population and Housing, and the responses to the 2006 Census Post Enumeration Survey for records that were successfully linked. Responses to the census question can be provided by the individual themselves or another member of the household. However, for the PES, only one adult is interviewed in the household to provide responses for all members of the household. Therefore it is possible for the person answering the questions for an individual may be different across the two collections.

Table 1. Summary of linked PES/Census records by Indigenous status , Australia–2006

	<u>PES response -</u>		
	Indigenous	Non-Indigenous	Total
Census response -	no.	no.	no.
Indigenous	2,576	283	2,859
Non-Indigenous	116	75,406	75,522
Not stated	38	1,218	1,256
Total	2,730	76,907	79,637

In respect of the births and deaths registrations, there are a similar range of issues, albeit to differing degrees. For births, which are sourced from the birth registration form submitted by parents and linked to hospital/midwife-completed notification to the Registrar for each birth, the general consensus is that Indigenous status of parents is well reported. Only a few jurisdictions request an Indigenous status for the baby as well as the parents and so Indigenous status is usually derived from that of the parents' response. Nevertheless, some volatility in the data, particularly at the sub-national level, indicates that some Indigenous births are not identified as such when they are registered. Furthermore there is evidence of considerable lags (approximately 5.7 months on average) between the Indigenous birth and the subsequent registration, compared to a national average of 2.3 months. To the extent that there is consistency in the lag in registrations this is not a great concern.

For deaths the issues are more serious. The identification of deaths as Indigenous or otherwise is dependent on appropriate identification on the death registration which is generally completed by the funeral director in consultation with a family member or other informant, or by a medical practitioner in the case of the separate medical certificate of cause of death. The process for completion, and the circumstances at the time of completion, makes the provision of accurate data a difficult goal. Contributing factors include non completion of the question, and completion without reference to any family members or other person who would have known the decedent's likely response. It would also be very insensitive and controversial to undertake any item non-response follow-up with the informant, or conduct a post enumeration survey to assess the response data quality. As a result the reporting of Indigenous status on death registrations is of variable quality, although there has been some improvement over time.

The data quality of the Census benchmark population estimates, and the components of intercensal and post-censal population growth, in particular the deaths, have precluded the compilation of annual population estimates of Indigenous Australians. Rather estimates of the population of Indigenous Australians are produced every 5 years based on the most recent census benchmarks and PES. An annual series is then compiled using assumptions for mortality and fertility, and reverse survival and projection techniques (ABS *Experimental Estimates and Projections, Aboriginal and Torres Strait Islander Australians, 1991 to 2021* (cat. No. 3238.0))

The Census Data Enhancement Project

In April 2005 the ABS issued *Discussion Paper Enhancing the Population Census: Developing a Longitudinal View* (cat. no. 2602.0) with a view to getting community reaction to a proposal to link census records over time, and to also link census data to a small number of administrative datasets including births and deaths registers for the purposes of statistical research.

Subsequent to the release of the discussion paper the ABS commissioned an independent Privacy Impact Assessment (PIA). The PIA made the following observation about the ABS proposal:

"For many people, the Proposal is likely to be seen as no more than a marginal change in the way in which the Census data is processed and used. No additional data is to be collected or held, and the only uses will be for statistical purposes within the well-established highly protective legislative and procedural framework, with no disclosure of identifiable personal information.

However, some will see the Proposal as a radical departure from established practice, which will create a data resource so rich and valuable for administrative uses that the privacy and secrecy framework under which the ABS operates may come under great and possibly irresistible pressure, if not immediately, then at least in the medium to long term."

The PIA also recommended that in respect of proposals involving name matching the ABS should be "confining the use of names during Census processing periods to ABS quality studies only".

Consequent to the many responses to the discussion paper and the PIA, the Australian Statistician issued a Statement of Intention on 18 August 2005 outlining a proposal to develop a Statistical Longitudinal Census Dataset based on a 5% random sample using statistical matching techniques, and to undertake a limited range of quality studies during the census processing period based on direct name/address matching. The ABS *Information Paper: Census Data Enhancement Project: An Update* (cat no 2602.0) provides more detail on the outcomes.

One of the quality studies approved was a quality study of Indigenous mortality data. The study would link data from the 2006 Census to death registrations for the period August 2006 to June 2007.

Indigenous Mortality Quality Study

As reported in *Information Paper: Census Data Enhancement – Indigenous Mortality Study* (cat no 4723.0) the CDE Indigenous Mortality Study involved linking Census records with death registration records to examine differences in the reporting of Indigenous status across the two datasets. The aims

of the project were to assess the under-identification of Indigenous deaths in death registration records and to identify factors that may be contributing to under-identification of Indigenous deaths.

As for all CDE studies the Indigenous Mortality Quality Study had strict security procedures in place to ensure confidentiality. The CDE Indigenous Mortality Quality Study linked 2006 Census records with deaths registered for the 11-month period following the Census. Death records examined related to deaths that occurred and were registered over the period. Records were linked during the Census processing period using name, address and other variables that were common across the two datasets, but excluding Indigenous status given that was the variable of interest. After Census processing was completed, all names and addresses held by the ABS on Census records and the linked dataset were destroyed. Following completion of analysis, the linked dataset was also destroyed.

Life Expectancy Estimates for Indigenous Australians

Due to the data quality issues around the five yearly Indigenous population estimates and the incomplete and varying coverage of Indigenous deaths from death registrations the compilation of Indigenous life expectancy estimates has been problematic.

Following the 1996 and 2001 Censuses the ABS compiled experimental life tables using different indirect demographic methods, (ie Preston and Hill method for 1991-1996 and Bhat method for 1996-2001) which were subject to a wide range of caveats.

Concerns about the sensitivity of indirect methods to assumptions and the opportunities presented by the above mentioned CDE quality study of Indigenous mortality data led to a further review of methods in 2008. These were presented in *Discussion Paper: Assessment of Methods for Developing Life Tables for Aboriginal and Torres Strait Islander Australians* (cat no 3302.0.55.002).

In particular, the discussion paper presented a direct demographic method using information from the CDE quality study of Indigenous mortality data to adjust recorded Indigenous deaths to a basis which was consistent with and comparable to the Census based population estimates. Specifically the CDE quality study of Indigenous mortality study showed the following in respect of the linked deaths data:

Table 2: Summary of linked deaths by Indigenous status, Australia–2006–2007

	<u>Death registration classification-</u>			
	Indigenous	Non-Indigenous	Not stated	Total
Census classification-	no.	no.	no.	no.
Indigenous	1,056	302	21	1,379
Non-Indigenous	231	91,076	970	92,277
Not stated	40	5,153	49	5,242
Total	1,327	96,531	1,040	98,898

In summary at the Australian level, of the 1,379 persons in the Census who identified as Indigenous, some 1,056 (77%) were identified as Indigenous in the death registrations. Countering this was the fact that 231 persons who identified as non-Indigenous in the Census identified as Indigenous in the death registration. The overall aggregate coverage was 96% (ie 1327/1379). From these data it was possible to develop 'identification rates and adjustment factors' at the state/territory level to apply to Indigenous death registrations. These were further adjusted to bring them in alignment to the PES so as to ensure consistency with the Indigenous population estimates. The availability of estimates of Indigenous deaths and populations on a consistent basis enabled direct methods to be applied to the compilation of Indigenous life expectancy estimates.

The final results of this analysis were presented in *Experimental Life Tables For Aboriginal and Torres Strait Islander Australians* (cat no 3302.0.5.003) which was released on 25 May 2009. The life expectancy at birth estimates for males was 67.2 years (with a 95% confidence interval between

65.8 and 68.6 years) and for females was 72.9 years (with a 95% confidence interval between 71.7 and 74.1 years). The difference in life expectancy between Indigenous and non-Indigenous Australians remains significant (11.5 years for males and 9.7 years for females) but is considerably less than the estimates based on the previous indirect methods (17 years for 1996-2001). For this reason the ABS has noted that the current (2005-2007) estimates are not comparable with the previous (1996-2001) Indigenous life expectancy estimates.

In discussion with key stakeholders and researchers, several alternative approaches were suggested and results prepared for comparison. Results ranged from 64.6 to 66.0 years for males and 70.6 to 71.9 for females (see Appendix 2 of *Experimental Life Tables for Aboriginal and Torres Strait Islander Australians* (cat no 3302.0.5.003))

The use of data linkage has considerably increased the capacity of the ABS to compile quality and robust Indigenous life expectancy estimates. ABS proposes to use the same methodology for compiling Indigenous Life Expectancy estimates for 2010-2012 to enable a comparable set of estimates over two census periods.

Conclusions

Developing good life tables and life expectancy estimates for Indigenous Australians remains a major challenge for the ABS. The use of data linking of administrative and population census records presents a significant opportunity to address some of the issues associated with the compilation of mortality statistics for Indigenous Australians.

However as demonstrated by the processes necessarily adopted by the ABS in the development of the Census Data Enhancement project the prospect of data linking presents significant privacy concerns amongst the community which need to be addressed. In addressing those concerns the ABS has decided to limit the scope of its CDE project.

Nevertheless, the availability of the CDE Quality Study of Indigenous Mortality Data has provided a very sound base on which to adjust recorded Indigenous deaths to a basis which is consistent with the Indigenous population estimates compiled from the Census and the PES. This in turn has provided a robust basis for the compilation of Indigenous life expectancy estimates that can inform on a critical policy issue in the Australian context.

By providing a direct estimate of under-identification of Indigenous deaths in death registrations, the ABS will be able to better work with the Registrars to further improve the quality of registrations. The ultimate goal, which may take years to realise, would be to publish annual estimates of the Indigenous population based on direct demographic methods along with associated life expectancies.

Australian Bureau of Statistics
September 2009

Bibliography

All of these Australian Bureau of Statistics releases are available on the ABS website
<http://www.abs.gov.au/>

ABS Experimental Estimates and Projections, Aboriginal and Torres Strait Islander Australians, 1991 to 2021 (cat. no. 3238.0)

ABS Experimental Life Tables for Aboriginal and Torres Strait Islander Australians (cat no 3302.0.55.003)

ABS Discussion Paper Enhancing the Population Census: Developing a Longitudinal View (cat. no. 2602.0)

ABS Information Paper: Census Data Enhancement Project: An Update (cat no 2602.0)

ABS Information Paper: Census Data Enhancement – Indigenous Mortality Study (cat no 4723.0)

ABS Discussion Paper: Assessment of Methods for Developing Life Tables for Aboriginal and Torres Strait Islander Australians (cat no 3302.0.55.002)

Acknowledgement

This paper is largely based on a similar ABS paper presented at the 57th Session of the International Statistics Institute, Durban, 16-22 August 2009 titled “The Challenge of Developing Good Quality Demographic Statistics for Indigenous Australians” authored by Peter Harper and Paul Sullivan.