



# Ethnic counts on mortality and census data 2001–06: New Zealand Census-Mortality Study update

Lavinia Tan, Tony Blakely, June Atkinson

### Abstract

**Aim** To provide an update for the assessment of discrepancies in ethnicity counts in the 2001 census and mortality data for the 2004–2006 period.

**Methods** 2001 census anonymously and probabilistically linked to 5 years of subsequent mortality data (135,849 eligible mortality records), allowing a comparison of ethnicity recording for the years 2001–2004 and 2004–2006.

**Results** Using a total definition of ethnicity, census and mortality counts agree reasonably well in 2004–06 and resemble comparisons in 2001–04, except at younger ages where counts for Pacific and Asian ethnicities are up to a third less for mortality data. Due to multiple ethnicities being more commonly recorded on census data, sole ethnicity counts are generally greater on mortality than census data, particularly for Māori ethnicity.

**Conclusion** Similar to 2001–2004, there is little bias in ethnic group counts between census and mortality data when using total ethnicity. Calculations of mortality rates by ethnicity using unlinked census and mortality data and a total definition of ethnicity should be unbiased. These results support ongoing use of the census definition of ethnicity on all health datasets.

The New Zealand Census-Mortality Study (NZCMS) has previously shown that Māori and Pacific deaths were significantly undercounted on mortality data relative to census data prior to 1995, and relatedly, non-Māori non-Pacific deaths overcounted on mortality data. In 1995, the ethnicity question on the death registration form was altered from the biological definition of race (blood more than half) to a more self-defined question resembling that on the 1996 census.

Thus, although historic mortality trends prior to 1995 required recalculation due to the considerable bias in ethnicity group counts, more recent NZCMS data has shown increasing agreement between census and 2001–04 mortality ethnic counts,<sup>1</sup> at least for the concept of "total ethnicity". This paper provides an update for the 2001 census linked to 2004–06 deaths, assessing any existing discrepancies between mortality and census data for ethnicity counts.

The main objectives of this paper are to examine any existing discrepancies in ethnicity counts for mortality (2004–06 and, by way of comparison, 2001–04) and 2001 census data, and how these differ depending on the ethnicity definition used. This paper provides an update for a series of publications focused on the numerator-denominator bias and changes in this bias over time.

# Methods

The methodology used in this paper, using on linked census and mortality data, has been described previously in technical reports and papers.<sup>2–7</sup> For the update 6 March 2004 to 5 March 2006 period, 79.8% of eligible mortality records were anonymously and probabilistically linked to census records. A subset of these records with highly probable links (HPL,76.1% of 2004–06 eligible mortality records), where ethnicity had no effect on linkage probability, was used in analyses to assess any discrepancy in ethnicity counts on the mortality and census data.

The number of deaths in the HPL dataset was weighted up so that it was representative of all 2004–06 eligible mortality records.<sup>7</sup> Weights were calculated based on variables that were predictors of HPL in logistic regression analyses: Age at census, sex, prioritised ethnicity, rurality, residential mobility of area unit, NZ deprivation index, Regional Health Authority, and cause of death. Cells within a stratum that met the numerical criterion of >3 linked records, were separated and assigned an independent weight, whereas the remaining cells were collapsed. The order of collapsing of strata variables to ensure sufficient cell sizes was based on the strength of their relationship with HPL.

Counts of ethnic groupings for census and mortality data were compared using weighted crossclassifications of the HPL dataset. Tabular count output is random rounded to a near multiple of three as per Statistics New Zealand protocol, but the census-mortality ratios were calculated on unrounded data.

### Ethnicity definitions

Three ethnicity definitions were used to assess any discrepancies in ethnic counts between the census and mortality data:

- *Total ethnicity* was assigned as Māori if any ethnic group identified on the census or mortality record was Māori. The same was done for Pacific, Asian and "non-Māori non-Pacific non-Asian" (nonMPA) ethnicities. Individuals could be assigned multiple ethnicities and consequently the sum of counts across ethnic groups will be greater than actual number of decedents.
- *Prioritised ethnicity* was assigned as Māori if any of the three possible self-identified ethnicity responses was Māori. For non-Māori, an individual was assigned as Pacific if one of the self-identified ethnic groups was Pacific. For non-Māori non-Pacific, an individual was assigned as Asian if one of the self-identified ethnic groups was Asian. The remaining individuals were assigned as nonMPA (equivalent to sole nonMPA).
- *Sole ethnicity* was assigned as Māori if Māori was the only ethnic group self identified. Similarly, an individual was assigned as Pacific or Asian if it was the only self-identified ethnic group. All others were assigned as "Remainder". This group also included some extra decedents that, for example, self-identified multiple ethnic groups.

# **Results**

Table 1 shows the weighted ethnicity counts for census and death registration form data, as well as the census to mortality ratios for the 2001–2004 and 2004–2006 cohorts for all three ethnicity definitions. There is generally close agreement between the census and mortality data for both cohorts in both the 2001–04 and 2004–06 data. There does appear to be some overcounting of the Māori ethnicities in mortality data relative to census data when using the sole ethnicity definition. This is due to fewer mortality records being assigned multiple ethnicity groups relative to the census data, resulting in greater sole Māori counts on the mortality data.

		2001-2	004+	2004-2006			
Ethnicity	Census	Mortality	Census to Mortality Ratio	Census	Mortality	Census to Mortality Ratio	
Total							
Māori	7,419	7,539	0.98	5,136	5,205	0.99	
Pacific	2,448	2,493	0.98	1,449	1,482	0.98	
Asian	1,236	1,215	1.02	804	816	0.99	
NonMPA‡	73,089	72,051	1.01	47,262	46,530	1.02	
Prioritised							
Māori	7,419	7,539	0.98	5,136	5,205	0.99	
Pacific	2,373	2,439	0.97	1,386	1,449	0.96	
Asian	1,170	1,155	1.01	777	771	1.01	
NonMPA‡	71,442	71,274	1.00	46,143	46,020	1.00	
Sole							
Māori	5,931	6,891	0.86	4,095	4,764	0.86	
Pacific	2,196	2,274	0.97	1,284	1,338	0.96	
Asian	1,098	1,086	1.01	744	735	1.01	
Remainder#	73,179	72,153	1.01	47,322	46,602	1.02	

# Table 1. Census and death registration form (mortality) ethnicity totals and ratios in 2001–2004 (n = 82,404 deaths), 2004–2006 (n = 53,445 deaths) for total, prioritised and sole ethnicity definition

Counts have been weighted and random rounded to a multiple of three as per Statistics New Zealand protocol. Note that the sum of observations for Prioritised and Sole ethnicities is equal to the total number of records, but the sum of observations for Total ethnicity is larger than this.

† The 2001–04 counts and ratios are sourced from Tables 10–13 of Fawcett et al. (2008)

<sup>‡</sup> The 'total nMnPnA' group was defined those people with one or more self-(undertaker-) defined ethnic groups, of which one was nMnPnA. The 'prioritized nMnPnA' is best thought of those remaining after all census respondents or decedents with any one of Māori, Pacific or Asian ethnicity have been 'prioritised out'. This is equivalent to the 'sole nMnPnA' group.

# The 'Remainder' group in sole ethnicity includes any people who reported nMnPnA ethnic group (i.e. the 'total nMnPnA' group) *plus* some extra decedents or census respondents who were recorded as, say, both Māori and Pacific and therefore not eligible to be either 'sole Māori' or 'sole Pacific'.

The counts for total ethnicity in 2001–04 and 2004–06, stratified by sex, age, regional health authority and NZ Deprivation Index are shown in Table 2. The counts for census and mortality data are largely congruent, and similar between the two cohorts The notable exception is the undercounting of total ethnicity in the mortality data for younger age groups (<25 years) for all ethnic groups, but especially for Pacific and Asian ethnicities. This is because multiple ethnicities are much more commonly identified on the census data than mortality data. Also note that the 2004–06 cohort counts are fewer and consequently less stable than the counts for the 2001–04 cohort.

Variables	Ethnicity		2001-2004		2004–2006			
		Census	Death Registration Form	Census to Mortality Ratio	Census	Death Registration Form	Census to Mortality Ratio	
Sex								
Male	Māori	4,095	4,140	0.99	2,793	2,841	0.98	
	Pacific	1,365	1,371	1.00	798	810	0.99	
	Asian	681	642	1.06	459	459	1.00	
	NonMPA	36,042	35,499	1.02	22,932	22,560	1.02	
Female	Māori	3,324	3,396	0.98	2,343	2,364	0.99	
	Pacific	1,086	1,122	0.97	651	669	0.97	
	Asian	552	573	0.97	348	357	0.97	
	NonMPA	37,047	36,552	1.01	24,333	23,970	1.02	
Age								
0–14 yrs	Māori	195	189	1.03	69	63	1.10	
	Pacific	63	60	1.02	21	15	1.40	
	Asian	18	21	0.91	18	12	1.50	
	NonMPA	354	309	1.14	126	105	1.20	
15–24 yrs	Māori	480	495	0.97	459	483	0.95	
	Pacific	174	156	1.10	132	111	1.19	
	Asian	87	75	1.16	39	27	1.44	
	NonMPA	1,161	1,062	1.10	1,017	912	1.12	
25-44 yrs	Māori	786	792	0.99	765	780	0.98	
	Pacific	237	234	1.01	186	180	1.03	
	Asian	114	105	1.09	87	96	0.91	
	NonMPA	2,064	1,947	1.06	1,890	1,779	1.06	
45–64 yrs	Māori	2,589	2,637	0.98	1,443	1,458	0.99	
	Pacific	732	726	1.01	363	375	0.97	
	Asian	282	261	1.09	171	168	1.02	
	NonMPA	9,348	9,048	1.03	5,157	4,989	1.03	
65–74 yrs	Māori	1,863	1,896	0.98	1,251	1,257	1.00	
	Pacific	558	579	0.97	339	345	0.98	
	Asian	282	285	0.98	192	201	0.96	
	NonMPA	12,852	12,648	1.02	7,512	7,383	1.02	
75–84 yrs	Māori	1,098	1,089	1.01	840	852	0.99	
	Pacific	480	522	0.92	294	321	0.92	
	Asian	297	297	1.00	189	198	0.95	
	NonMPA	23,937	23,775	1.01	15,765	15,636	1.01	
85+ yrs	Māori	411	441	0.93	315	306	1.03	
	Pacific	204	213	0.96	120	138	0.87	
	Asian	156	174	0.90	111	114	0.97	

Table 2. Census and death registration form Total ethnicity counts and ratios by sex, age, RHA and NZ Deprivation in 2001–04 (n=82,404), 2004–06 (n= 53,445)

NZMJ 13 August 2010, Vol 123 No 1320; ISSN 1175 8716 URL: <u>http://www.nzma.org.nz/journal/123-1320/4269/</u>

Variables	Ethnicity	2001–2004			2004–2006			
		Census	Death Registration Form	Census to Mortality Ratio	Census	Death Registration Form	Census to Mortality Ratio	
	NonMPA	23,370	23,262	1.00	15,804	15,726	1.00	
RHA								
Northern	Māori	2,253	2,298	0.98	1,566	1,596	0.98	
	Pacific	1,716	1,758	0.98	1,023	1,059	0.97	
	Asian	684	687	0.99	459	471	0.97	
	NonMPA	19,902	19,578	1.02	12,681	12,486	1.02	
Midland	Māori	2,835	2,913	0.97	1,965	1,989	0.99	
	Pacific	174	177	0.99	102	84	1.21	
	Asian	129	111	1.17	63	69	0.91	
	NonMPA	14,850	14,454	1.03	9,831	9,591	1.03	
Central	Māori	1,713	1,743	0.98	1,200	1,227	0.98	
	Pacific	459	444	1.03	246	255	0.96	
	Asian	276	279	0.99	198	189	1.05	
	NonMPA	19,080	18,891	1.01	12,390	12,186	1.02	
Southern	Māori	618	585	1.05	405	393	1.03	
	Pacific	99	114	0.88	78	84	0.93	
	Asian	144	135	1.06	90	90	1.00	
	NonMPA	19,257	19,125	1.01	12,357	12,267	1.01	
NZ Depriv	ation							
1–4	Māori	312	303	1.03	696	672	1.04	
	Pacific	84	93	0.89	147	174	0.84	
	Asian	273	267	1.02	303	318	0.95	
	NonMPA	11,781	11,709	1.01	16,977	16,824	1.01	
5-6	Māori	597	603	0.99	687	678	1.01	
	Pacific	159	174	0.93	183	189	0.97	
	Asian	237	234	1.01	180	171	1.05	
	NonMPA	14,565	11,460	1.01	11,118	11,022	1.01	
7–8	Māori	1,062	1,011	1.05	1,140	1,164	0.98	
	Pacific	273	294	0.93	276	282	0.98	
	Asian	231	228	1.00	171	177	0.97	
	NonMPA	16,794	16,659	1.01	10,992	10,839	1.01	
9–10	Māori	1,659	1,704	0.97	2,616	2,691	0.97	
	Pacific	486	477	1.02	840	834	1.01	
	Asian	282	282	1.00	150	153	0.98	
	NonMPA	17,451	17,163	1.02	8,175	7,845	1.04	

All counts are weighted, and then random rounded to a multiple of three as per Statistics New Zealand protocol. Minimum cell size is 6.

Full cross-classifications of census and mortality 2004–2006 data for prioritised and sole ethnicity definitions are shown in Table 3 and 4. These show the specific mismatches in ethnicity counts between the census and mortality files. The majority

of misclassifications were between Māori and nonMPA groups. Using a sole definition of ethnicity, the number of Māori deaths according to mortality data exceeds census data (i.e. census-mortality ratio <1.0), due to multiple self-identified ethnicities being more common on census data, this causes sole Māori counts to be lower on census than mortality data.

Sex	Census	Māori	Pacific	Asian	NonMPA	Total Census	Census to
	Prioritised	Deaths	Deaths	Deaths	Deaths	Deaths	<b>Mortality Ratio</b>
	Ethnicity						
Males	Māori	2,556	9	6	222	2,793	0.98
	Pacific	9	723	6	27	756	0.96
	Asian		18	405	18	444	1.02
	Non MPA	276	39	30	22,023	22,368	1.00
	Total	2,841	789	435	22,296		
Females	Māori	2,133	6		204	2,343	0.99
	Pacific	12	594	6	21	630	0.95
	Asian		15	312	9	336	1.01
	Non MPA	219	45	18	23,487	23,772	1.00
	Total	2,364	660	333	23,724		

Table 3. Census by mortality counts for prioritised ethnicity by sex, 2004–06

All counts are weighted, and then random rounded to a multiple of three as per Statistics New Zealand protocol. Minimum cell size is 6.

Sex	Census Sole Ethnicity	Māori Deaths	Pacific Deaths	Asian Deaths	Remainder Deaths	Total Census Deaths	Census to Mortality Ratio
Males	Māori	2,394	6	6	390	2,268	0.87
	Pacific	9	693	6	60	714	0.98
	Asian		6	396	42	417	0.98
	Non MPA	204	27	24	22,110	22,965	1.02
	Total	2,604	732	426	22,602		
Females	Māori	1,992	6		345	1,827	0.85
	Pacific	9	570	6	48	570	0.94
	Asian		6	294	39	324	1.04
	Remainder	159	33	12	23,568	24,357	1.01
	Total	2,160	609	312	24,003		

### Table 4. Census by mortality counts for sole ethnicity by sex, 2004–06

All counts are weighted, and then random rounded to a multiple of three as per Statistics New Zealand protocol. Minimum cell size is 6.

# Discussion

Using a total definition of ethnicity, census and mortality counts agree reasonably well in 2004–06 and were similar to 2001–04, except at younger ages where mortality data counts for Pacific and Asian are up to a third (1-1.5) less for mortality data – although numbers are sparse and hence estimates unstable. Due to multiple ethnicities being more commonly recorded on census data, prioritised and sole comparisons differ; Māori sole counts are greater on mortality than census data.

The current analyses necessarily used only a subgroup of the total mortality records (the highly probable links, HPL) to estimate numerator denominator bias. By weighting up these HPL links to be representative of all mortality records, we are assuming that within strata of that weighting (i.e. sex by age by prioritised ethnicity (mortality data), rurality, residential mobility of area unit, NZ deprivation index, region, and cause of death) that the (dis)agreement of mortality and census data ethnicity is the same among the HPL dataset as the non-HPL dataset. Unfortunately, we cannot prove this assumption. However, we are reasonably confident that within all cross-classifications of these strata that we essentially adjust for any selection bias that may arise in using the subsidiary HPL data-set.

The use of a total definition of ethnicity appears to be the most accurate and consequently if this definition is used for the calculation of ethnic mortality rates, there should be little to no numerator-denominator bias. As with our last update, these results again provide support to the sector for ongoing attempts to ensure health data uses an ethnicity question as close in wording and layout to the census question as possible.

### Competing interests: None known.

Author information: Lavinia Tan, Research Fellow; Tony Blakely, Research Professor; June Atkinson, Data Manager/Senior Data Analyst; Department of Public Health, Wellington School of Medicine and Health Sciences, University of Otago, Wellington

Acknowledgments: The NZCMS is conducted in collaboration with Statistics New Zealand and within the confines of the Statistics Act 1975. The NZCMS was funded by the Health Research Council of New Zealand, and is now funded by the Ministry of Health.

Access to the data used in this study was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. The results presented in this study are the work of the authors, not Statistics New Zealand.

**Correspondence:** Tony Blakely, Department of Public Health, Wellington School of Medicine and Health Sciences, Health Inequalities Research Programme, University of Otago, Wellington, PO Box 7343, Wellington, New Zealand. Fax: +64 (0)4 3895319; email: tony.blakely@otago.ac.nz

### **References:**

- Blakely T, Atkinson J, Fawcett J. Ethnic counts on mortality and census data (mostly) agree for 2001-2004: New Zealand Census-Mortality Study update. NZ Med J 2008; 121(1281):58-62.
- Ajwani S, Blakely T, Robson B, et al. Unlocking the numerator-denominator bias for the 1980s and 1990s. NZCMS Technical Report No. 4. ISBN 0-473-09111-9. Wellington: Department of Public Health, Wellington School of Medicine and Health Sciences, University of Otago, 2002.
- 3. Ajwani S, Blakely T, Robson B, et al. Unlocking the numerator-denominator bias III: adjustment ratios by ethnicity for 1981-1999 mortality data. The New Zealand Census-Mortality Study. NZ Med J 2004; 116 (1175):6.

- 4. Blakely T, Atkinson J. Unlocking the Numerator-Denominator Bias, 1991-94 Deaths. NZCMS Technical Report No. 2. Wellington: Department of Public Health, Wellington School of Medicine, University of Otago, 2001.
- Hill S, Atkinson J, Blakely T. Anonymous record linkage of census and mortality records: 1981, 1986, 1991, 1996 census cohorts. NZCMS Technical Report No. 3. ISBN 0-473-09110. Wellington: Department of Public Health, Wellington School of Medicine and Health Sciences, University of Otago, 2002.
- Fawcett J, Atkinson J, Blakely T. Record linkage of census and mortality 2001-04 records. New Zealand Census-Mortality Study Technical Report No. 6. Wellington: Health Inequalities Research Programme, University of Otago, Wellington, 2008.
- 7. Tan L, Blakely T, Atkinson J. Linkage bias and numerator-denominator bias in census and mortality 2001-06 records. NZCMS Technical Report No 7. Wellington: Health Inequalities Research Programme, University of Otago, Wellington, 2010.