Official Statistics Research Series, Vol 5, 2009
ISSN 1177-5017; ISBN 978-0-478-31593-6

The Official Statistics System

Does Mortality Vary Between Pacific Groups?

Estimating Samoan, Cook Island Mäori, Tongan, and Niuean mortality rates using hierarchical Bayesian modelling

Tony Blakely
Health Inequalities Research Programme, University of Otago

Ken Richardson
Health Inequalities Research Programme, University of Otago

Jim Young
Vital Statistics Ltd

Patrick Graham
Department of General Practice and Public Health, University of Otago

Martin Tobias
Ministry of Health

Paul Callister
Victoria University of Wellington

Robert Didham
Statistics New Zealand

June Atkinson
Health Inequalities Research Programme, University of Otago

Citation

Blakely, T., Richardson, K., Young, J., Graham, P., Tobias, M., Callister, P., Didham, R., & Atkinson, J. (2009). Does mortality vary between Pacific groups? Estimating Samoan, Cook Island Māori, Tongan and Niuean mortality rates using hierarchical Bayesian modelling, *Official Statistics Research Series, 5.*

Available from http://www.statisphere.govt.nz/official-statistics-research.aspx

Published by

Statistics New Zealand
Tatauranga Aotearoa
Wellington, New Zealand

ISSN 1177-5017 (online)
ISBN 978-0-478-31593-6 (online)

Abstract

Pacific mortality rates are traditionally presented for all Pacific people combined, yet there is likely heterogeneity between separate Pacific groups. We aimed to determine mortality rates for Samoan, Cook Island Māori, Tongan and Niuean people. We also aimed to test and demonstrate the application of hierarchical Bayesian methods to sparse data. We used New Zealand Census-Mortality Study (NZCMS) data for 2001–04, for 380,000 person years of follow-up of 0–74 year olds in the 2001–04 cohort for which there was complete data on sex, age, ethnicity (total counts), country of birth and household income. Given sparse data, we used hierarchical Bayesian (HB) regression modelling, with: a prior covariate structure specified for sex, age, country of birth (CoB) and household income; and smoothing of rates using shrinkage. The posterior mortality rate estimates were then directly standardised. Standardising for sex, age, income and CoB, all-cause mortality rate ratios compared to Samoan were: 1.21 (95 percent credibility interval 1.05 to 1.42) for Cook Island Māori; 0.93 (0.77 to 1.10) for Tongan; and 1.07 (0.88 to 1.29) for Niuean. Cardiovascular disease (CVD) mortality rate ratios showed greater heterogeneity: 1.66 (1.26 to 2.13) for Cook Island Māori; 1.11 (0.72 to 1.58) for Niuean; and 0.86 (0.58 to 1.20) for Tongan. Results were little different standardising for just sex and age. We conducted a range of sensitivity analyses about a plausible range of (differential) return migration by Pacific people when terminally ill, and a plausible range of census undercounting of Pacific people. Our findings, in particular the elevated CVD mortality among Cook Island Māori, appeared robust. To our knowledge, this project is the first time in New Zealand that clear (and marked in the case of CVD) differences in mortality have been demonstrated between separate Pacific ethnic groups. Future health research and policy should, wherever possible and practicable, evaluate and incorporate heterogeneity of health status among Pacific people.

Keywords

Pacific Islanders, mortality rates, hierarchical Bayesian modelling.

Acknowledgements

Tony Blakely conceived and lead study overall, led the writing, and contributed to study design, interpretation.

Ken Richardson led the hierarchical Bayesian analyses, and contributed to the overall study design, interpretation, and writing.Jim Young and Patrick Graham provided advice and technical assistance for the hierarchical Bayesian analyses, and contributed to the overall interpretation of study findings.

Martin Tobias provided advice, and provided particular input to sensitivity analyses.

Paul Callister provided advice, and particular input to the writing of the demographic components of the Introduction.

Robert Didham was the Statistics New Zealand sponsor of the project. He also provided advice, and particular input to the writing of the demographic components of the Introduction.

June Atkinson conducted the standardisation analyses, and provided data and technical assistance and advice.

Contents

Tables

Figures

Executive summary

Background

Due to relatively small numbers Pacific mortality rates are traditionally presented for all Pacific people combined, not by specific Pacific groups (eg, Niuean). Yet there is good reason to believe that considerable heterogeneity in mortality rates exist between separate Pacific peoples. This project aimed to determine mortality rates during 2001–04 for Samoan, Cook Island Māori, Tongan and Niuean people who were usually resident in New Zealand on 2001 census night.

Methods

We used aggregated and confidentialised unit record census-mortality data from the New Zealand Census-Mortality Study (NZCMS). Most analyses were based on (i) total definitions of ethnicity, ie, if an individual reported multiple ethnicities, and one of those was Samoan, then that individual was counted as being Samoan, and (ii) 380,000 person years of follow-up of 0–74 year olds in the 2001–04 cohort for which there was complete data on sex, age, ethnicity, country of birth (New Zealand versus other) and income. Thirty eight percent of 0–74 year old person time was excluded due to missing data, principally missing income.

We initially calculated directly standardised mortality rates as routinely done for official statistics, and in most previous NZCMS analyses.

We then used hierarchical Bayesian (HB) regression modelling, with prior covariate structure specified for sex, age, country of birth and household income. Smoothing was also incorporated in the modelling by allowing posterior mortality rates to move closer to that actually observed in a given stratum as data richness in that stratum increased, using shrinkage estimation. HB modelling was undertaken for models adjusting for just sex and age, additionally for country of birth, and additionally for household income. The posterior mortality rate estimates from these models were then directly standardised to the total Pacific 2001–04 population as above to allow direct comparability.

We conducted a range of basic sensitivity analyses about likely residual systematic biases (eg, return to home country migration when terminally ill).

Finally, we also trialled a methodological extension to HB methods, namely allowing the global fit prior shrinkage parameter to vary according to local data.

Results

HB modelling – substantive findings

Adjusting just for sex and age, 0–74 year old all-cause mortality rates in 2001–04 were estimated as:

- Cook Island 288 per 100,000 (95 percent credible intervals 255 to 325). Rate ratio compared to the largest group, Samoan, was 1.18 (1.02 to 1.37)
- Niuean 252 per 100,000 (208 to 297). Rate ratio compared to Samoan 1.03 (0.84 to 1.24)
- Samoan 244 per 100,000 (224 to 266).
- Tongan 234 per 100,000 (197 to 269). Rate ratio compared to Samoan 0.96 (0.80 to 1.12).

The rates changed by less than 3 percent when additionally adjusting (and standardising) for country of birth and household income. However, as the Cook Island rate tended to increase further and the Tongan rate decreased further, rate ratios compared to Samoan in the completely adjusted analysis for Cook Island was 1.21 (1.05 to 1.42) and for Tongan was 0.93 (0.77 to 1.10) – a 30 percent higher mortality rate for Cook Island Māori compared to Tongan.

By cause of death, differences in the fully adjusted mortality rates among Pacific groups were most notable for cardiovascular disease (CVD):

- Cook Island 111 per 100,000 (90 to 135). Rate ratio compared to Samoan 1.66 (1.26 to 2.13).
- Niuean 75 per 100,000 (51 to 104). Rate ratio compared to Samoan 1.11 (0.72 to 1.58).
- Samoan 68 per 100,000 (61 to 84).
- Tongan 58 per 100,000 (40 to 78). Rate ratio compared to Samoan 0.86 (0.58 to 1.20).

That is, 0–74 year old CVD mortality was nearly twice as high among Cook Island Māori compared to Tongan.

Moderating these strong CVD differences were weaker (and perhaps opposing) non-statistically significant differences in cancer. For injury/suicide mortality combined, Niuean people had a rate ratio of 0.53 (0.17 to 0.99) compared to Samoan people. Injury/suicide mortality rates for Samoan, Cook Island Māori and Tongan were similar.

HB modelling compared to routine direct standardisation

The pattern of estimates, and the magnitude of estimates, was similar for both routine direct standardisation and standardisation after HB modelling. Thus, whilst HB modelling has strong theoretical advantages, there was no substantial differences between empirical and HB standardised results for the highly aggregated statistics reported here. Nevertheless, the replication of results by HB modelling increases confidence in the results of routine direct standardisation. The real strength of the HB approach would most likely be seen by comparing results for data-poor strata, such as mortality rates for young males.

HB modelling – methodological findings

Allowing the prior model for zeta (the parameter affecting shrinkage of stratum-specific mortality rates to the prior model) to vary according to the amount of locally adjacent data did not notably alter HB modelling estimates.

Sensitivity analyses

Using the posterior mortality rates estimates from HB modelling, sensitivity analyses suggested that the elevated Cook Island Māori mortality rate (especially that for CVD mortality) is robust. These analyses considered a plausible range of (differential) return migration by Pacific people when terminally ill, and a plausible range of census undercounting of Pacific people.

Conclusions

HB modelling

We successfully implemented HB modelling on linked census-mortality data to assess Pacific group-specific mortality rate variations. Undertaking this HB modelling was time consuming and, therefore, costly in terms of human resource. Significant researcher up-skilling was required, highly detailed cross-classified data had to be extracted from the Statistics New Zealand Data Laboratory, and then time consuming (human and CPU time) and complex HB modelling had to be undertaken off-site from the Data Laboratory. For future projects, some of these barriers will be reduced – especially the research-team skills.

We have shown that HB models can be developed for sparse-data problems, and on the basis of this work we should be able to answer questions not easily approached using more traditional methods. For example, using this model we can predict mortality rates across strata of age, sex, and country-of-birth. It might be possible to go further and address the question of what is driving the high mortality rates observed for Cook Island Māori people. The ‘sparse-data’ problem is, by definition, an area of interest to official statistics research, but we note that Bayesian approaches are also being applied to the analysis of longitudinal data, missing and misclassified data, imputation, complex survey design, and the generation of synthetic datasets. We anticipate that this trend will accelerate in the future both within New Zealand and internationally.

Pacific mortality

To our knowledge, this project is the first time in New Zealand that clear (and marked in the case of CVD) differences in mortality have been demonstrated between Pacific people. This is extremely important for further health research and policy. At present, policy approaches to Pacific health disparities often assume that they are the same across ethnic subgroups. Our results demonstrate that this is not the case, and support other (emerging) evidence of diversity in health status within the Pacific population. Researchers should, if at all possible, try and analyse and present results separately for specific ethnic groups within the Pacific grouping, and attempt to explain the causes and contexts of these differences. Policymakers increasingly need to consider Pacific group-specific policies and programmes.

1 Introduction

1.1 Demography of Pacific people

Pacific mortality cannot be considered without some understanding of the migration history of Pacific people to New Zealand. As part of this history, Te Ara (2008) notes that the migration of Pacific peoples is linked to colonial histories.[A] [B] They note that Tahitians and New Caledonians have tended to migrate to France, American Samoans to America, and Melanesians to Australia. The Pacific peoples who have migrated to New Zealand have come mainly from nearby islands, and from those countries with a British colonial history.

Māori are the consequence of the first Pacific migration to New Zealand which arrived over a period of time starting perhaps 1,000 years ago.(King 2003) In the early period of European colonisation small numbers of Pacific people lived in New Zealand. For example, Te Ara (2008) states that the 1872 New Zealand census recorded 31 people born in the 'South Sea Islands'. By 1916, 18 Melanesians, 49 Fijians and 151 'other and undefined' Polynesians had settled in New Zealand.

As the need for more unskilled labour increased in the 1960s and 1970s, immigration from some areas of the Pacific was sought after to fill labour shortages. Cook Islanders and Niueans since 1901 and Tokelauans since 1916 had already become New Zealand citizens.

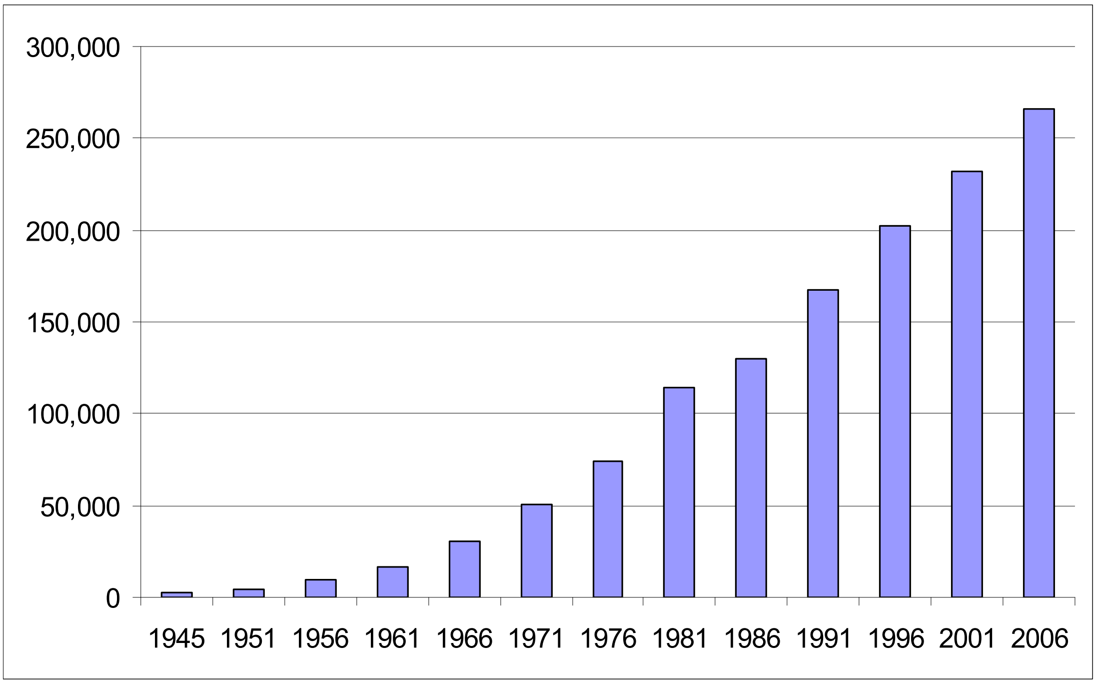
A Treaty of Friendship was signed with the Samoan government in 1962, and the Western Samoan Quota scheme was established to facilitate migration from Samoa. The Quota provided residence to 1,100 Western Samoan citizens annually conditional on having definite employment in New Zealand (irrespective of skill or labour market needs), age (18–45 years) and standard health and character requirements. In 1973 a major review of immigration policy was undertaken. One outcome of the review was that in 1974 immigration policy reaffirmed the free access to New Zealand of those born in the Cook Islands, Niue, and Tokelau. A 'Pacific Access Category' also provides limited access to migrants from Tonga, Kiribati, Tuvalu, and Fiji.

As a result of labour demand, changes in migration policy, and humanitarian responses following major storms in Tokelau, a large number of migrants from the Pacific Islands, mainly Samoa, Tonga, the Cook Islands, Niue and Tokelau arrived in New Zealand during the late 1950s to 1970s to provide a valuable source of labour in the urban and manufacturing sector. In 1945 the Pacific population was just over 2,000 people, Samoans being the largest group. However since the 1960s the Pacific population had been rapidly increasing. In recent times, the population was 202,233 in 1996, rising to 231,801 in 2001 and increasing further to 265,974 in March 2006 (Callister and Didham 2008).

[A] South Pacific peoples, Te Ara, http://www.teara.govt.nz/NewZealanders/NewZealandPeoples/SouthPacificPeoples/1/en

[B] There is increasing literature on Pacific migration to New Zealand (eg Cook, L., Didham, R., & Khawaja, M. (1999). On the demography of Pacific people in New Zealand. Wellington: Statistics New Zealand., Fairburn-Dunlop, P., & Makisi, G. S. (2003). Making our place: growing up PI in New Zealand. Palmerston North: Dunmore Press., Anae, M., Iuli, L., & Burgoyne, L. (2006). The Polynesian Panthers 1971-1974: the crucible years. Auckland: ReedMacpherson, C. (2004). From Pacific Islanders to Pacific People: The Past, Present and Future of the Pacific Population in Aotearoa. In P. Spoonley, C. Macpherson & D. Pearson (Eds.), Tangata, Tangata: The Changing Contours of Ethnicity in Aotearoa New Zealand. Wellington: Thomson Dunmore Press— (2006). Pacific People in Aotearoa/New Zealand: from Sojourn to Settlement. In K. Ferro & M. Wallner (Eds.), Migration Happens. Reasons, Effects and Opportunities of Migration in the South Pacific. Vienna: Lit Verlag GmbH & Co. KG.

Figure 1
Pacific population in New Zealand, 1945 to 2006



Source: Statistics New Zealand

The six largest groups of Pacific peoples in New Zealand are Samoans, Cook Islanders, Tongans, Niueans, Fijians and Tokelauans. However, living in New Zealand are Austral Islanders, Belau/Palau Islanders, Bougainvilleans, Caroline Islanders, Easter Islanders, Gambier Islanders, Guam Islanders, Hawaiians, I-Kiribati, Kanaka and other New Caledonians, Marquesas Islanders, Marshall Islanders, Nauruans, Papua New Guineans, Phoenix Islanders, Pitcairn Islanders, Tahitians, Solomon Islanders, Tuamotu Islanders, Tuvaluans, Vanuatuans, Wallis Islanders and Yap Islanders.

Table 1 shows the census based populations of the six largest Pacific groups living in New Zealand and the estimated resident populations of the countries associated with these ethnicities. A number of points stand out. In relation to this mortality study, for both Cook Island Māori and Nuieans, the population in New Zealand is substantially higher than the ‘home’ population (4.14 and 14.05 times greater, respectively). In contrast the ‘home’ Samoan population is larger than the New Zealand population. But caution is needed with these data. As Bedford (Bedford 2008) points out, there is considerable movement between New Zealand and these islands, especially for Niue and Cook Islands where the number of people who at some point over a year lived on the home islands is much higher.

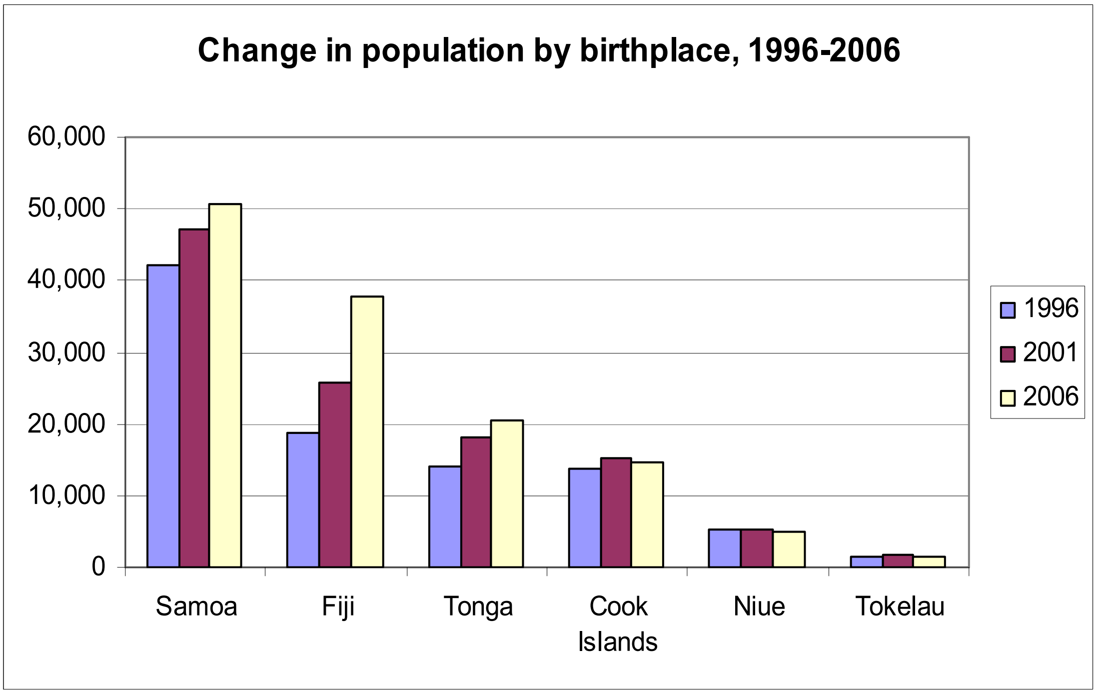
Table 1
Size of main Pacific ethnic groups, 1996 to 2006, Total counts, and estimated resident population in home country

	1996	2001	2006	Estimated resident population home country	Ratio of number living in NZ compared to number living in home country
Samoan	101,754	115,026	131,103	182,700	0.72
Cook Island Mäori	47,019	51,486	58,008	14,000	4.14
Tongan	31,392	40,716	50,481	98,300	0.51
Niuean	18,477	20,154	22,476	1,600	14.05
Fijian	7,695	7,041	9,861	917,000	0.01
Tokelauan	4,917	6,198	6,822	1,392	4.90

(1) Source of New Zealand data, Statistics New Zealand
(2) The data for Tokelau are from https://www.cia.gov/cia/publications/factbook/countrylisting.html, while the remaining data are taken from the Pacific Forum website http://www.forumsec.org/pages.cfm/about-us/member-countries/. Some data are for 2004 and some for 2006. Approximately 51 percent of the population of Fiji (2007 estimate) are ethnic Fijians.

Figure 2 switches the focus to changes in population by country of birth. It demonstrates that the Samoan population is also the largest when Pacific people are restricted to those born in the Pacific.

Figure 2
Pacific population born in Pacific Island, but usually resident in New Zealand – trends from 1996 to 2006



Source: Statistics New Zealand

The numbers in Figure 2 are significantly smaller than in Table 1, but 60 percent of Pacific people living in New Zealand were born in New Zealand. The proportion of the Pacific population born in New Zealand has been steadily increasing. In 1976, 38 percent were born in New Zealand. By 1991 this had reached 50 percent and by 2006 60 percent. But as Table 2 shows there are differences in the proportion born in New Zealand when specific groups are looked at. The ethnic group with the highest proportion born in New Zealand is Niueans at 74 percent.

Table 2
Percentage of each main Pacific group who were born in New Zealand, 2006, total Pacific group counts

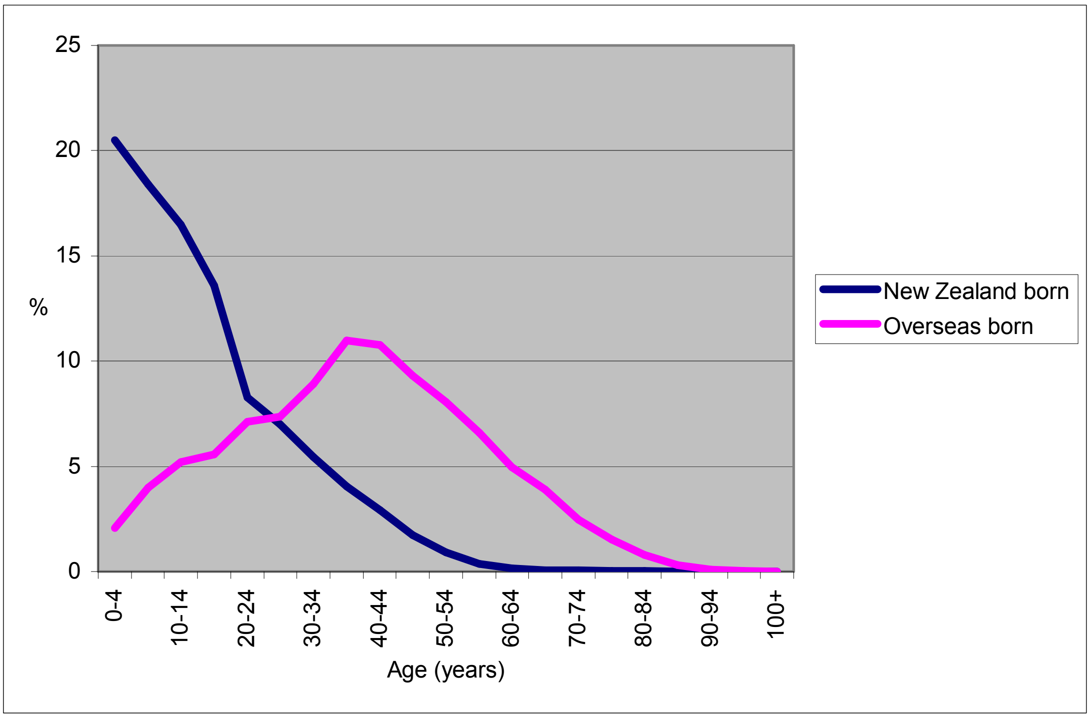
	New Zealand born (percent)
Niuean	74.1
Cook Island Mäori	73.4
Samoan	59.7
Tongan	56.0
Total Pacific ethnic group	60.0

Source: Statistics New Zealand

The Pacific population living in New Zealand is young, particularly those of its members born in New Zealand. Figure 3 shows there are marked differences in age structures between those Pacific people born in New Zealand and those born outside New Zealand.[C] Just under 70 percent of the New Zealand born Pacific population was under 20 as against just 17 percent of those born overseas. However, it is worth noting that there are strong connections between these populations. In many households there will be parents or grandparents who were born overseas and children or grandchildren born in New Zealand, and links are maintained across the communities and with family living elsewhere in New Zealand, as well as in other countries.

[C] A small number of those born outside New Zealand will also been born outside of Pacific countries. Examples will include those Pacific people born in Australia.

Figure 3
Age distribution of the New Zealand born and overseas born New Zealand Pacific population, 2006



Source: Statistics New Zealand

Table 3shows a range of socio-demographic factors by Samoan, Niuean, Cook Island Māori and Tongan, according to 2006 and/or 2001 census data. Notable differences include:

- Tongan and Cook Island Māori have the youngest median age.
- Whilst all Pacific groups have at least 50 percent of their population now born in New Zealand, Cook Island Māori and Niuean have the highest percentage born in New Zealand.
- Educational attainment is highest for Samoan, and least for Cook Island Māori.
- Of particular note for health, the prevalence of current smoking among adults at the 2001 census was greatest for Cook Island Māori (38 percent), least for Samoan (28 percent) and Tongan (29 percent), and intermediate for Niuean (33 percent).

Table 3
Some demographic data for Samoan, Cook Island Māori, Tongan, and Niuean at the 2006 census (and 2001 census for selected variables)

Variable	Census year	Samoan	Cook Island Māori	Tongan	Niuean
Number	2006	131,103	58,011	50,478	22,476
	2001	115,000	52,600	40,700	20,100
Percent of total Pacific population	2006	49%	22%	19%	8%
	2001	50%	23%	18%	9%
Sole any-Pacific † #	2006	77%	66%	82%	68%
Sole group-specific Pacific ‡ #	2006	66%	53%	71%	41%
Median age (years) #	2006	20.9	18.9	18.6	19.6
Current smokers #	2001	28%	38%	29%	33%
Percent of adults with formal Qualification #	2006	69%	55%	64%	60%
Percent living in Auckland #	2006	67%	60%	80%	79%
Born in NZ	2006	60%	73%	56%	74%
	2001	58%	70%	53%	70%
Duration of residence (a)	2006	6%	18%	6%	18%
Percent speaking own language	2006	63%	16%	61%	25%
Percent reporting religion	2006	86%	70%	90%	70%
Percent living in extended family	2006	35%	32%	39%	35%
Median personal income for adults	2006	$21,400	$19,800	$17,500	$21,500
Own home (b)	2006	23%	21%	19%	21%
Access to telephone	2006	83%	76%	81%	79%

(a) Percent of those born overseas residing in New Zealand for less than one year.

(b) Percent of adults in fully or partly owned home.

† 'Sole any-Pacific' means the person did not self-identify as any non-Pacific groups, but might have self identified as two or more specific Pacific groups (eg, Samoan and Niuean).

‡ 'Sole specific-Pacific' means the person only identified one Pacific ethnic group – and nothing else (eg, Samoan only).

Only 2006 reported due to either missing data for 2001 or little change from 2001.

Source of much of the 2006 data, Pacific Profiles: 2006, http://www.stats.govt.nz/analytical-reports/pacific-profiles-2006/default.htm

Tables 4 and 5 focus on the main populations within the wider Pacific People group. Using the more detailed level 3 ethnic groups (Statistics New Zealand level 3; eg, Sri Lankan and Indian categorised separately, and Pacific groups categorised as in this project), the two tables explore how likely is it that a person will form a partnership with someone from their own ethnic group, such as a Samoan marrying a Samoan (these are total count data, so the person may have also recorded other ethnicities as well). The tables also show, given that the person does not have a partner from the same level 3 ethnic group, how likely it is that their partner will also be from the wider Pacific ethnic group.

Tongans and Samoans are the most likely to have a partner from the same ethnic group as themselves. But for Tongans and Samoans, if a person from these groups does not have a partner from the same level 3 ethnic group, it is more likely that the partner will have an ethnicity from outside of the Pacific group than within it. There are also gender differences underlying these data. For instance, Tongan and Samoan women are more likely to have a partner from their same ethnic group than are men (ie, men are less likely to have a partner from their own group).

These data indicate some considerable variation in ethnic intermarriage rates within the wider Pacific ethnic group. But the intermarriage data also indicate there is some social distance between the subgroups within the wider Pacific Peoples ethnic group, for example some distance between Samoans and Tongans.

Table 4
Partners of Pacific women from main groups – Ranked by whether their partner is from the same level 3 ethnic group, opposite sex couples, Total counts, 2006

	Males						
	Same ethnic group#	**Other Pacific**	**European**	**Mäori**	**Asian**	**MELAA**	**Total couples**
Tongan female	**74.6**	8.8	12.8	5.0	1.3	0.1	6,711
Samoan female	**70.4**	5.9	18.5	7.2	2.2	0.3	18,150
Cook Island Mäori female	**40.9**	15.3	28.5	20.6	2.0	0.4	6,825
Niuean female	**38.1**	26.5	25.8	14.7	2.7	0.5	2,535

MELAA = Middle East, Latin America, Africa.
# But the partner may have also recorded other ethnic groups as well.
Source: Statistics New Zealand

Table 5
Partners of Pacific men from main groups – Ranked by whether their partner is from the same level 3 ethnic group, opposite sex couples, Total counts, 2006

	Females						
	Same ethnic group#	**Other Pacific**	**European**	**Māori**	**Asian**	**MELAA**	**Total couples**
Tongan male	**67.6**	11.0	15.8	11.3	1.2	0.2	7,407
Samoan male	**65.5**	6.5	23.1	12.5	2.2	0.2	19,512
Cook Island Māori male	**41.2**	8.5	34.0	29.4	1.8	0.2	6,780
Niuean male	**35.2**	21.2	31.8	23.6	2.7	0.1	2,748

MELAA = Middle East, Latin America, Africa.
# But the partner may have also recorded other ethnic groups as well.
Source: Statistics New Zealand

One of the outcomes of ethnic intermarriage is more complex ethnicities among their children. Using the data of the 42,160 Pacific children for the period 2000–2004, we can see that over half (54 percent) of all Pacific children have at least one other ethnicity. While over half (53 percent) of the births report only Pacific ethnicities (22,605 out of 42,160), 30 percent report Māori as their ethnicity and 27 percent report at least one European ethnicity. These are, of course, not additive, with nearly half (48 percent) of the Pacific/Māori children also having a European ethnicity.[D]

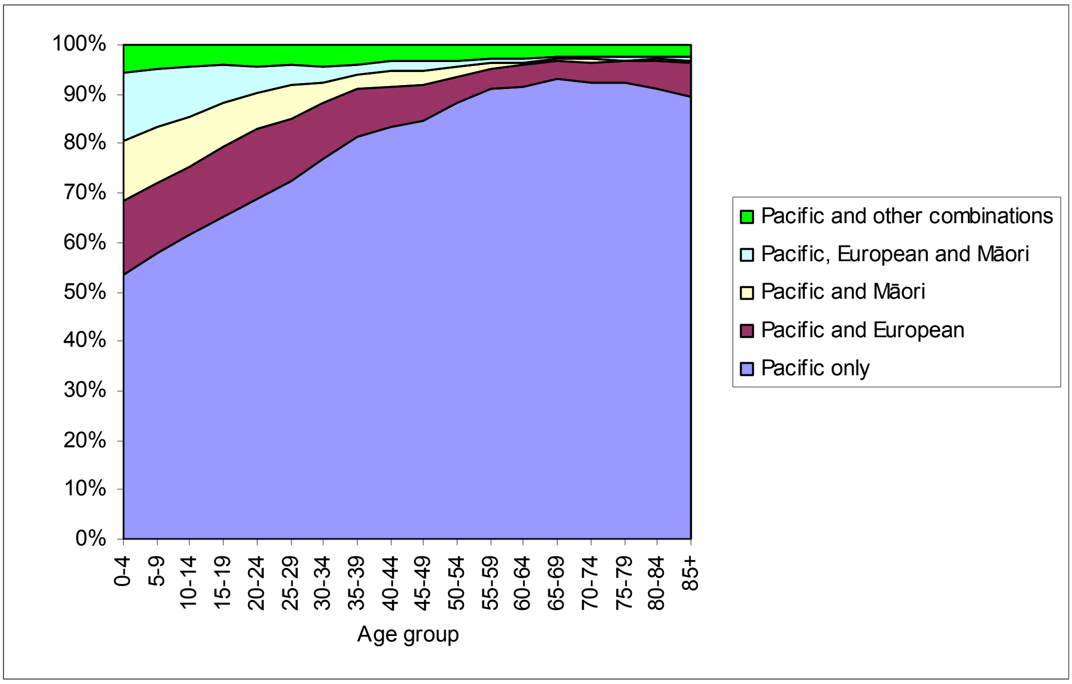
Table 6
Ethnicities of Pacific children born in main groups 2000–2004

	Number of births	Percent with more than one ethnicity	Percent with more than one Pacific ethnicity
Samoan	21,194	53.6	19.4
Cook Island	9,890	72.0	25.5
Tongan	9,624	45.8	23.5
Niuean	3,722	84.8	48.6
Total Pacific	42,160	54.0	23.0

Source: Statistics New Zealand

While it is far more common amongst young people to have multiple ethnicities reported in official surveys, figure 4 shows that in all age groups there are Pacific people who record more than one ethnic group. Particularly common amongst young Pacific people is a combination with Māori, either as one of two combinations or part of the combination Pacific/Māori/European.

Figure 4
Main single and combination ethnic responses for Pacific people by age, 2006



Source: Statistics New Zealand

1.2 Health variations between Pacific groups

Health statistics usually present rates of risk factors, disease and mortality for all Pacific people combined (Blakely et al 2007; Ministry of Health 2008; Ministry of Health and Ministry

[D] In the past, dual and multiple ethnicities amongst Pacific people led to some undercount of the Pacific population (see Appendix).

of Pacific Island Affairs 2004). Yet, Pacific people are heterogeneous in culture, socioeconomic and health status.

Sundborn et al (2008) have recently published estimates of a range of risk factors for Samoan, Tongan, Niuean and Cook Island Māori participating in the Diabetes Heart and Health Study (DHHS) (Sundborn et al 2008a). These are reproduced in **Error! Reference source not found.** and **Error! Reference source not found.** of the Appendix. Being a workforce survey, the population is considerably older than the census population. The smoking prevalence also differs from census data, raising concerns about representativeness of the DHHS. However, it must be noted that the DHHS results are presented by sex, and there is substantial variation in smoking behaviour among Pacific people by sex. If one takes the average of men and women smoking from the DHHS: Samoan people have the highest smoking prevalence (26.1 percent – they had the lowest prevalence according to census data); and Niuean have the lowest smoking prevalence (15.6 percent – they had an intermediary prevalence according to census data). There may be very plausible reasons why risk factor patterns in the DHHS may differ from the census (and hence presumably the Pacific population in New Zealand), including healthy worker effect, age structure, geographical location, etc. Nevertheless, the DHHS suggests substantial variation in health risk factors by Pacific group.

1.3 Objectives of this project

This Official Statistics Research (OSR) funded research project has the following objectives:

1. Is it feasible to estimate reliable smoothed mortality rates using hierarchical Bayesian methods for Samoan, Cook Island Māori, Tongan and Niuean people? Including:
    a. differences in causes of death for major disease groupings (eg, CVD, cancer)
    b. differences in mortality rates between sole and combination groupings of Pacific people (eg, [Samoan] compared to [Samoan and European])
2. Explore methodological extensions to hierarchical Bayesian methods, specifically allowing second level parameters that determine *global* fit of data to the prior to vary *locally* depending on adjacent data.

There is growing interest in HB methods in empirical social sciences, especially epidemiology. The ability to incorporate prior information, and to incorporate smoothing (or more technically shrinkage), is particularly attractive in light of the issue at hand – determining mortality rates for small populations with sparse data. They give some protection against model mis-specification while nevertheless permitting some smoothing of crude rates. Hierarchical models have been shown to outperform classical regression in predictive accuracy (Gelman 2006). Bayesian 'shrinkage' estimators (discussed below) have good variance reduction properties, particularly when sample size is small (Best et al 2005; Greenland 2008).

The following sections describe the methods and data used in this paper – and the hierarchical Bayesian methods in detail. Next, we present mortality rates using 'routine' direct standardisation methods. The HB results are then given using models that adjust for (1) just sex and age, and (2) more fully for sex, age, country of birth and household income. This is followed by basic sensitivity analyses for possible biases affecting the calculation of Pacific group-specific mortality rates. Finally, we conclude with some 'lessons learnt' and recommendations.

2 Methods

2.1 Data

Analyses were conducted at the University of Otago, Wellington after aggregation and confidentialisation of unit record 2001–04 NZCMS data available in the Statistics New Zealand Data Laboratory. Details of the linkage, weighting for incomplete linkage of mortality data to census data, and variables specifications can be found elsewhere (Fawcett et al 2008).

For the majority of analyses, we restricted the dataset to person-years of follow-up for people aged 0–74 years with non-missing data on ethnicity, sex, age, and equivalised household income. These restrictions were made for the following reasons:

- Household income and country of birth were required for final models to assess the contribution of these variables to Pacific differences.
- The upper age limit was set at 74 years because first, whilst the 2001–04 cohort included people of all ages, previous cohorts only include people up to age 74 on census night; and second, household income becomes less predictive of mortality among retired people.
- Household income and country of birth were required for final models.

The distribution of data (census counts, unweighted counts of linked mortality records, and weighted counts of linked mortality records) are shown by sex, age-group, and Pacific group in **Error! Reference source not found.**.

Table 7
Census and death counts by sex by age, 2001-04 cohort, 'total' definition of ethnicity

		Total Pacific			Total Samoan			Total Cook			Total Tongan			Total Niuean		
Sex	Age at End FU	Census Count	Linked Deaths	Deaths (weighted)	Census Count	Linked Deaths	Deaths (weighted)	Census Count	Linked Deaths	Deaths (weighted)	Census Count	Linked Deaths	Deaths (weighted)	Census Count	Linked Deaths	Deaths (weighted)
M&F	All ages	215,319	1590	2172	107,316	744	1020	48,372	384	519	38,097	237	321	18,747	144	201
	0–14 yrs	70,713	45	60	35,391	21	27	17,133	12	15	14,064	12	15	6,438	6	6
	15–34 yrs	74,316	138	216	37,272	69	105	16,713	33	51	12,465	27	39	6,495	9	12
	35–44 yrs	29,835	120	168	14,607	54	72	6,306	36	51	5,046	18	27	2,439	6	9
	45–64 yrs	31,263	516	693	15,831	252	339	6,261	132	174	4,911	66	84	2,460	45	60
	65–74 yrs	6,156	372	495	2,877	171	228	1,368	90	114	1,014	45	60	594	45	60
	75+	3,033	396	546	1,332	180	246	591	81	114	594	69	96	318	42	57
Males	All ages	105,243	897	1230	52,437	402	555	23,685	222	303	18,921	147	198	9,222	81	111
	0–14 yrs	36,192	30	39	18,117	12	18	8,832	6	9	7,170	9	12	3,273	6	6
	15–34 yrs	35,925	90	147	17,916	45	75	8,115	24	36	6,102	21	30	3,222	6	9
	35–44 yrs	13,890	66	96	6,804	27	39	2,856	21	27	2,457	15	18	1,140	6	6
	45–64 yrs	15,219	300	408	7,791	144	198	2,985	81	108	2,466	36	51	1,197	27	33
	65–74 yrs	2,799	216	285	1,317	93	126	642	54	72	453	24	33	282	30	39
	75+	1,218	192	258	495	78	99	252	39	51	270	45	60	114	15	24
Females	All ages	110,076	690	942	54,876	339	462	24,687	159	216	19,176	90	123	9,525	66	90
	0–14 yrs	34,521	18	21	17,274	6	9	8,301	6	6	6,897	6	6	3,168	6	6
	15–34 yrs	38,391	45	66	19,356	21	30	8,598	9	18	6,360	6	9	3,273	6	6
	35–44 yrs	15,945	54	69	7,803	24	30	3,450	15	21	2,592	6	9	1,302	6	6
	45–64 yrs	16,044	216	285	8,043	105	141	3,273	54	69	2,445	27	36	1,263	18	24
	65–74 yrs	3,360	156	210	1,560	78	102	723	33	42	561	21	30	315	15	18
	75+	1,815	204	291	843	102	150	339	42	60	324	27	36	207	27	36

Note: All counts are random rounded to a near multiple of three as per Statistics NZ protocols. Minimum cell size is 6.

Sex was classified as a dichotomous variable, with the reference group being male. Age was categorised in five year age groups for the direct standardisation analyses. For the HB modelling five age-groups were used: 0–14 years, 15–34 years, 35–44 years, 45–64, and 65–74 years, centred at the 35–44 age group, and scaled so that each unit increase in scaled age corresponds to an actual increase of 10 years. To allow for the non-linear increase in mortality with age, a linear spline with knots at the 35–44 and 45–64 age groups was fitted. Thus the above age ranges are represented by their end-points which, after centring and scaling become elements from the set {-3, -1, 0, 2, 3}. Age is assumed to vary linearly with the model, with discontinuities in the slope at 35–44 (0) and 45–64 (2). It should be noted that the age stratification might be improved in future work by disaggregating 45–64 into two ten-year age groups.

The ethnicity variable was classified in two ways. First, it was categorised using a total definition. Thus, all of the following people would be categorised as Niuean:

- self-identified Niuean only
- self-identified Niuean and Samoan
- self-identified Niuean and New Zealand European.

This total classification meant that some people were classified in two (or perhaps more) of the four Pacific groups used in the study: Samoan; Cook Island Māori; Tongan; Niuean. This total classification was used for the standardisation and majority of HB modelling. We also used a sole classification (eg, Samoan only, with no other Pacific or wider ethnic group affiliation) for some additional HB models to assess the impact of the method of classification on patterns between the four Pacific groups.

Country of birth (CoB) was simply coded as ‘New Zealand born’ versus ‘non-New Zealand born’, the latter being overwhelmingly equivalent to born in the country of their stated ethnicity.

We used equivalised household income, categorised into tertiles within strata of sex and age-group, and based on income for the New Zealand Pacific population. Income ranks were median centred and scaled (divided by 10). Thus, income ranks (coded as 1, 2, 3) are transformed to (-0.1, 0, 0.1) and variation of income was assumed to be log linear.

The standardisation analyses were conducted in the Data Laboratory at Statistics New Zealand. Bayesian modelling was conducted at the University of Otago, Wellington, using aggregated estimates of the weighted un-standardised death rates and person-years for each possible combination of sex (2) by age group (5) by Pacific ethnic group (4) by country-of-birth (2) by income group (3). That is, a total of 240 cells, each containing a mortality rate, number of person years of follow up, and (by multiplying the former two) a number of observed deaths. A relatively coarse stratification of age and income (referred to as the restricted dataset in below) was used for the Bayesian analysis to reduce the impact of small data cells as much as possible. Posterior mortality rates for each stratum were computed by standardising to the New Zealand Pacific population.

2.2 Standardised rates

Directly standardised mortality rates were calculated using the same restricted dataset as for the HB modelling, and using methods applied frequently before to NZCMS data (eg, Blakely et al 2007). The total 2001 census Pacific population was used as the direct standard for weights by age (in five-year age groups). Note that the most comparable HB analyses also used age standardisation, but additionally standardised for sex. The absence of standardisation for sex in the routine standardisation analyses is not expected to change the results substantially.

2.3 Modelling – hierarchical Bayesian Poisson regression

We largely follow the HB methods used previously in the NZCMS. Young et al (2006) applied the hierarchical Bayesian regression approach of Christiansen and Morris (1997) for Poisson count and rates using linked census-mortality data. In brief their method was as follows. Assuming death is a Poisson process such that for Pacific ethnicity $j\,(=1,\cdots,4)$ and stratum $i\,(=1,\cdots,60)$ with deaths d_{ij}, mortality rate λ_{ij}, and person-years at risk P_{ij}, a three-level Poisson model was defined by:

$$d_{ij} \mid \lambda_{ij}, P_{ij} \sim \text{Poisson}\,[\lambda_{ij} P_{ij}, \lambda_{ij} P_{ij}]\,, \tag{1}$$

$$\lambda_{ij} \mid X_i, \beta_j, \varsigma \sim \text{gamma}[\mu_{ij}, \mu_{ij} / \varsigma_j^{\,2}]\,, \tag{2}$$

$$\log(\mu_{ij}) = X_i \beta_j\,, \tag{3}$$

$$\beta_j, \varsigma_j \sim \pi\,. \tag{4}$$

The mortality rate λ_{ij} had a gamma distribution with mean μ_{ij} and variance μ_{ij}/ζ_j^2, and the prior mean μ_{ij} had a structure that depended on covariates X_i and parameters β_j through a log-link function. Second-level parameters, β_j (the regression 'hyper-parameters') and ζ_j (the mortality rate variance or 'shape' hyper-parameter) were assigned independent prior distributions ('hyper-priors') at the third level of the hierarchy. The generic distributions π are described below.

Extending the Young et al model to allow for variation by Pacific ethnicity, the regression hyper-parameter vector was partitioned as $\beta_j = (\beta_{1j}, \beta_2)$ to allow some of the components (the β_{1j}) to vary by Pacific ethnicity (in this study, the intercept; see further discussion below). A standard approach was adopted for β_j, with uniform prior distributions for each component.

The prior covariate structure influences the mean of the posterior rate, but the degree of influence depends on the overall support for the prior covariate structure in the data, as well as on how much local information is available. Given the structure of the model defined by equations (1) and (2), the conditional posterior distribution for the mortality rate is also gamma with mean

$$\mathrm{E}[\lambda_{ij} \mid \mathbf{y}, \beta_j, \varsigma] = B_{ij}\mu_{ij} + (1 - B_{ij}) y_{ij}\,, \tag{5}$$

where $y_{ij} = d_{ij} / P_{ij}$ is the observed mortality rate in the ith stratum of the jth ethnicity, $\mathbf{y} = (y_1, y_2, \ldots)$ and

$$B_{ij} = \varsigma_j / (\varsigma_j + \mu_{ij} P_{ij})\,. \tag{6}$$

Thus, the conditional mean for λij is a weighted average of the prior mean μij and the observed mortality rate (yij). The Bij, which lie between zero and one, are known as shrinkages because larger values shrink the conditional posterior mean mortality rates towards the prior mean. The gamma shape parameter ζj provides a measure of in the influence of the prior mean – large values of ζj correspond to shrinkages close to one. The relatively uninformative (Daniels 1999) uniform shrinkage prior of Christiansen and Morris was adopted for ζj: Following equation (6), defining $B_{0j} = \varsigma_j / (\varsigma_j + \varsigma_{0j}) \sim uniform(0,1)$ yields a prior distribution for ζj with median ζ0j. Two choices for ζ0j are investigated below. In the first case, ζ0j is fixed at 10 for every Pacific ethnic group so that the shrinkage prior is the same for every ethnicity stratum (ie, $\zeta_{0j}=\zeta_0$). This concept of fit is global, referring to the fit of the prior model to the dataset as a whole. Given that data sparseness is a central issue for this analysis, ζ_{0j} was also allowed to vary across Pacific ethnicity, permitting greater shrinkage to the prior mean in data-poor strata. The potential advantage of this more local concept of fit approach is that it could prevent real ethnic group differences from being swamped by the mortality pattern of the majority group, while still permitting smoothing via shrinkage towards the prior model for the mortality rates.

A priori, we expected interaction of age and income, and sex and income as predictors of the mortality rate. Thus the two components of the regression hyper-parameters in equation (3) for the most complex prior model were

$$\begin{aligned} \beta_{1j} &= \beta_{j,0} \\ \beta_2 &= (\beta_{sex}, \beta_{age}, \beta_{CoB}, \beta_{inc}, \beta_{age\times inc}, \beta_{sex\times inc}). \end{aligned} \tag{7}$$

Prior models adjusted for just sex and age, additionally for country of birth, and additionally for household income. Models with income also included the interaction terms described above.

We exploit the Bayesian hierarchical approach by specifying a (prior) covariate model. Two prior models were studied. In the first we assumed mortality rate differences between Samoan, Tongan, Cook Island and Niuean people to be similar across, say, age and sex and imposed a prior covariate structure that did not depend on ethnic groups. However, the DHSS results suggest that variation of the prior covariate structure across Pacific groups might also be reasonable. We investigate this alternative hypothesis by allowing the intercept ($\beta_{j,0}$) in (7) to vary by ethnicity (j). Imposing such prior covariate models helps reduce instabilities that arise due to small numbers of observations in each cell. On the other hand, because of the overall hierarchical model structure, we are not forcing posterior values to conform exactly to the prior covariate model. When $\beta_{j,0}$ (= β_0) does not depend on ethnic group, differences in posterior conditional estimates occur only because they are shrunk towards the empirical means as in equation (5).

We used posterior mortality rate estimates from the models to calculate directly standardised rates and rate ratios, with 95 percent credible intervals (CIs). The distribution of the total eligible Pacific person-time was used as the standard population, ie, for the fully-stratified analysis, 60 weights formed by cross-classifying the sex by five age groups (ie,ie, 0–14, 15–34, 35–44, 45–64, and 65–74 years) by income (tertiles) by CoB (2) strata of person-time.

Thus, total person-time in each age, sex, income, and CoB (*a*,*s*,*t,b*) stratum was computed by summing over (total) ethnicities $P_{\bullet astb} = \sum_j P_{jastb}$. Weights were derived by dividing total stratum person-time by total Pacific person-time $P = \sum_{astb} P_{\bullet astb}$, ie, $w_{astb} = P_{\bullet astb} / P$.

Standardisation over strata of <u>a</u>ge, <u>s</u>ex, income <u>t</u>ertile, and country of <u>b</u>irth (denoted ASTB) for ethnicity *j* was calculated from posterior estimates of stratum-specific mortality rates in the usual way: $\lambda^*_{\text{j,ASTB}} = \sum_{astb} \lambda_{\text{jastb}} w_{astb}$. For the partially-stratified case (eg, by age and sex only), deaths and person-time were aggregated across unwanted strata, weights and posterior mortality rates calculated for the remaining strata, and standardisation proceeded in a similar manner to that described above, eg, $\lambda^*_{\text{j,AS}} = \sum_{as} \lambda_{\text{jas}} w_{as}$. To enable consistent comparison of results across Bayesian analyses, standardisation weights based on total ethnicities were used in all cases – including those that used the sole definition of ethnicity.

As noted above, the main results reported here are based on data that were fully cross-classified by Pacific ethnicity, sex, age, and income, with all main effects included and with interaction terms for income with age and sex, and standardised by sex, age, income and CoB when calculating overall Pacific mortality rates. It was also of interest to see the impact on ethnic rates of not adjusting for income (an intermediary between ethnicity and mortality – not a confounder) or CoB. For these extra models, it was necessary to aggregate the input data across levels of the variable omitted from modelling.

For each of the models we ran seven parallel chains with 120,000 iterations. The first 60,000 iterations of each chain were discarded, and each chain was thinned by retaining every 60th value. This resulted in a posterior sample size of 7,000 available for inference. Convergence was assessed using the Gelman-Rubin statistic (Brooks and Gelman 1998) which was close to one for all parameters.

Model comparisons were made, where appropriate, using the deviance information criterion Spiegelhalter et al (2002) which was available in the version of WinBUGS (1.4) used here. Gelman et al (2004) suggest that this criterion is useful when the goal is choosing a model with the best out-of-sample predictive power. Graphical plots of posterior predictive standardise rates were used to check models by assessing differences between real and simulated data (Brooks and Gelman 1998).

Software

All analyses and plots were done using the R environment (http://www.r-project.org) for statistical computation version 2.2.0 available from the Comprehensive R archive Network (CRAN) website (http://cran.r-project.org) or SAS 8.2 (SAS Institute Inc., Cary, North Carolina). All HB analyses used WinBugs 1.4, available from (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml), and the R bugs function from the CRAN package R2WinBUGS version 2.0-4.

2.4 Sensitivity analyses

We examined the likely impact of two possible biases:

- return migration
- census under-enumeration.

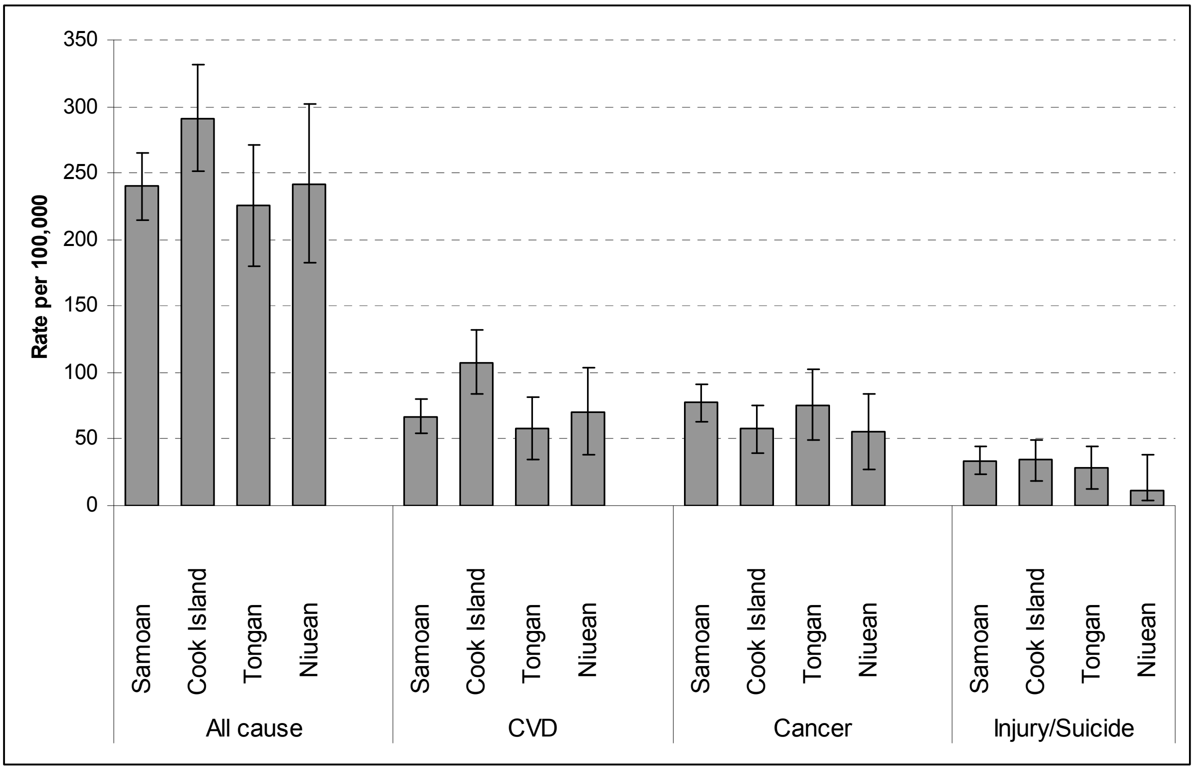
Methods are straightforward and best described at the same time as results are presented.

3 Results

3.1 Directly standardised rates

All-cause and cause specific (CVD, cancer, and injury/suicide) age-standardised mortality rates per 100,000 are show in Figure 5, Table 8, and Table 9.

Figure 5
Age-standardised mortality rates (per 100,000) for 0–74 year olds, sexes combined, 2001–04, using 'routine' methods – total Pacific groups, same restricted data-set as used for HB analyses



Cook Island Māori had the highest all-cause mortality rate and Tongan the lowest (although 95 percent confidence intervals overlapped).

This ranking was reproduced again, albeit more starkly, for CVD with the rate for Cook Island Māori (107.8 per 100,000) being 87 percent higher than for Tongan (57.8), with non-overlapping 95 percent confidence intervals. The CVD rates for Niuean and Samoan were intermediate, although closer to the Tongan rate, meaning that the Cook Island Māori CVD rate 'stands out' as particularly high.Cancer rates showed a different pattern, with lower rates for Cook Island Māori and Niuean – but 95 percent confidence intervals all substantially overlapped. The Niuean injury/suicide rate was particularly low, but with very wide 95 percent CI that largely overlapped the somewhat similar rates for the three other Pacific groups.

Table 8
Age-standardised all-cause mortality rates (per 100,000) and rate ratios (compared to Samoan) for 0–74 year olds in 2001–04, by sex. Same restricted data-set as used in HB modeling

Sex	Pacific group	Person years	Deaths (weighted)	Rate	(95% CI)	Rate ratio
Combined	All Pacific	384,106	1008	248.9	(230.9-266.9)	
	Samoan	190,924	474	239.8	(214.4-265.2)	1
	Cook Island	92,925	273	291.4	(250.9-331.8)	1.22
	Tongan	60,679	132	225.8	(180.3-271.4)	0.94
	Niuean	35,215	87	242.0	(182.3-301.7)	1.01
Males	All Pacific	189,862	621	309.8	(281.0-338.6)	
	Samoan	94,443	285	288.6	(248.8-328.4)	1
	Cook Island	45,939	174	380.3	(313.4-447.3)	1.32
	Tongan	30,564	84	283.0	(211.2-354.8)	0.98
	Niuean	17,547	57	322.5	(225.0-419.9)	1.12
Females	All Pacific	194,244	387	189.0	(167.2-210.8)	
	Samoan	96,481	186	191.3	(159.6-223.1)	1
	Cook Island	46,986	96	206.1	(159.3-253.0)	1.08
	Tongan	30,115	45	164.8	(109.8-219.9)	0.86
	Niuean	17,668	30	164.2	(93.5-234.9)	0.86

(1) Error bars are 95 percent CI. Ethnic groups are ‘total’ ethnic groups.
(2) The standard population is the total 2001–04 Pacific population, ie, excluding those with missing data, and those not at their usual residence nor at a private dwelling on census night.

Table 9
Age-standardised cause-specific mortality rates (per 100,000) and rate ratios (compared to Samoan) for 0–74 year olds in 2001–04, sexes combined. Same restricted data-set as used in HB modeling

Cause of death	Pacific group	Person years	Deaths (weighted)	Rate	(95% CI)	Rate ratio
CVD	All Pacific	384,106	303	74.1	(64.5-83.6)	
	Samoan	190,924	132	67.1	(54.0-80.3)	1
	Cook Island	92,925	99	107.8	(83.8-131.8)	1.61
	Tongan	60,679	33	57.8	(34.8-80.8)	0.86
	Niuean	35,215	24	70.5	(37.6-103.5)	1.05
Cancer	All Pacific	384,106	288	70.4	(60.8-80.0)	
	Samoan	190,924	153	77.3	(62.9-91.8)	1
	Cook Island	92,925	54	57.4	(39.1-75.7)	0.74
	Tongan	60,679	45	75.7	(49.4-102.1)	0.98
	Niuean	35,215	21	55.4	(27.1-83.6)	0.72
Injury/	All Pacific	384,106	117	30.3	(23.6-37.0)	
Suicide	Samoan	190,924	69	33.7	(23.5-43.8)	1
	Cook Island	92,925	30	34.1	(19.0-49.2)	1.01
	Tongan	60,679	18	27.9	(11.8-44.0)	0.83
	Niuean	35,215	6	11.6	(3.6-37.7)	0.34

(1) Error bars are 95 percent CI. Ethnic groups are ‘total’ ethnic groups.
(2) The standard population is the total 2001–04 Pacific population, ie, excluding those with missing data, and those neither at their usual residence nor at a private dwelling on census night.

Table 10 shows age-standardised all-cause mortality rates, and rate ratios compared to Samoan, for the four total Pacific groups for:

- 0–74 year olds in 2001–04 using the restricted dataset used in HB analyses presented later in this report (ie, as presented above too)
- 0–74 year olds in 2001–04 using all data (ie, including those with missing income)
- all ages in 2001–04 using the full data-set, and
- 0–74 year olds in 1996–99 using full data-set.

Two conclusions are possible:

- Cook Island Māori consistently have the highest mortality rate
- Tongan consistently have the lowest mortality rate in 2001–04, but Samoan have the lowest rate in 1996–99 – although confidence intervals between Samoan, Niuean, and Tongan substantially overlap.

Table 10
Age-standardised all-cause mortality rates (per 100,000) and rate ratios (compared to Samoan), sexes combined, by data-set restrictions, and for 1996–99 and 2001–04

		2001–04			1996–99
		Restricted dataset	Full dataset	Full dataset	Full dataset
Pacific group	**Measure**	**0–74 yrs**	**0–74 yrs**	**All ages**	**0–74 yrs**
All Pacific	Rates	248.9	244.1	319.9	295.7
Samoan		239.8	235.9	315.3	275.5
Cook Island		291.4	289.3	368.2	359.4
Tongan		225.8	207.8	275.2	299.2
Niuean		242.0	248.0	318.2	315.0
All Pacific	Rate ratios				
Samoan		1	1	1	1
Cook Island		1.22	1.23	1.17	1.30
Tongan		0.94	0.88	0.87	1.09
Niuean		1.01	1.05	1.01	1.14

(1) Ethnic groups are 'total' ethnic groups. The standard population is the 2001–04 total Pacific population.

Table 11 shows the CVD mortality rate for the same restrictions on datasets and for 1996–99. Cook Island Māori consistently have the highest CVD mortality rate, on average being 50 percent greater than for Samoan.

Table 11
Age-standardised CVD mortality rates (per 100,000) and rate ratios (compared to Samoan), sexes combined, by data-set restrictions and 1996–99 and 2001–04

		2001–04			1996–94
		Restricted dataset	Full dataset	Full dataset	Full dataset
Pacific group	Measure	0–74 yrs	0–74 yrs	All ages	0–74 yrs
All Pacific	Rates	74.1	74.9	109.2	99.6
Samoan		67.1	69.3	102.2	90.6
Cook Island		107.8	110.5	153.3	127.5
Tongan		57.8	55.5	86.0	108.9
Niuean		70.5	68.1	101.5	97.1
All Pacific	Rate ratios				
Samoan		1	1	1	1
Cook Island		1.61	1.59	1.50	1.41
Tongan		0.86	0.80	0.84	1.20
Niuean		1.05	0.98	0.99	1.07

(1) Ethnic groups are 'total' ethnic groups. The standard population is the 2001–04 total Pacific population.

3.2 Hierarchical Bayesian modelling

3.2.1 Allowing ζ (zeta) to vary made no substantive difference

We tested the impact of allowing the prior shrinkage parameter ζ to vary by Pacific ethnicity using 'Total' ethnicity input data, a prior model that adjusted for age, sex, CoB, and income (and a priori interactions) and posterior rates standardised by the same variables. An otherwise equivalent model that forced a common prior shrinkage parameter across ethnicity strata was used for the comparison. For this simpler model the prior distribution for ζ (from equation (6) is determined by $B_0 = \varsigma/(\varsigma+10) \sim uniform(0,1)$ – for more details see Section 0. The simpler model generated a substantially smaller deviance information criterion (846.5 vs 866.8 for the more complex case), and very similar parameters for the prior model (see

Table 12). The standard deviations, however, tend to be less when the zeta prior can vary. On balance, however, there is little support for allowing the shrinkage parameter prior to vary by ethnicity, and results reported below all refer to uniform shrinkage priors that are the same across ethnicity. It is possible that shrinkage priors may vary across other gradients of data scarcity, such as age and sex, but we do not pursue these ideas here.

Table 12
Posterior estimates of prior model parameters for zeta prior being common across ethnicity strata, and where it varies with ethnicity

	zeta prior common		zeta prior varies	
parameter	value	sd	value	sd
β_{0C}	-5.69	0.14	-5.65	0.12
β_{0N}	-5.82	0.14	-5.80	0.14
β_{0S}	-5.81	0.12	-5.78	0.11
β_{0T}	-5.87	0.14	-5.88	0.12
β_{inc}	-0.08	0.04	-0.07	0.03
β_{sexF}	-0.63	0.09	-0.62	0.08
β_{age1}	0.73	0.08	0.74	0.07
β_{age2}	-0.06	0.12	-0.10	0.11
β_{age3}	0.60	0.14	0.62	0.12
β_{NZbrn}	-0.26	0.11	-0.27	0.10

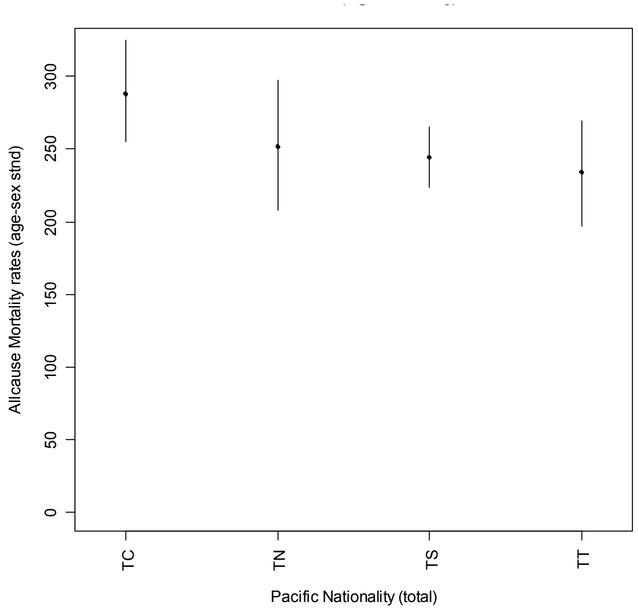
3.2.2 All-cause, 'Total' ethnicity

Adjusting for and standardising by sex and age

Posterior standardised rates for all-cause mortality from a prior model that included only age and sex and an intercept that varied by ethnicity, and which was standardised for the same variables is shown in Figure 6 and **Error! Reference source not found.** (Model 1). As noted above, the underlying data were stratified by age, sex, and Pacific ethnicity only.

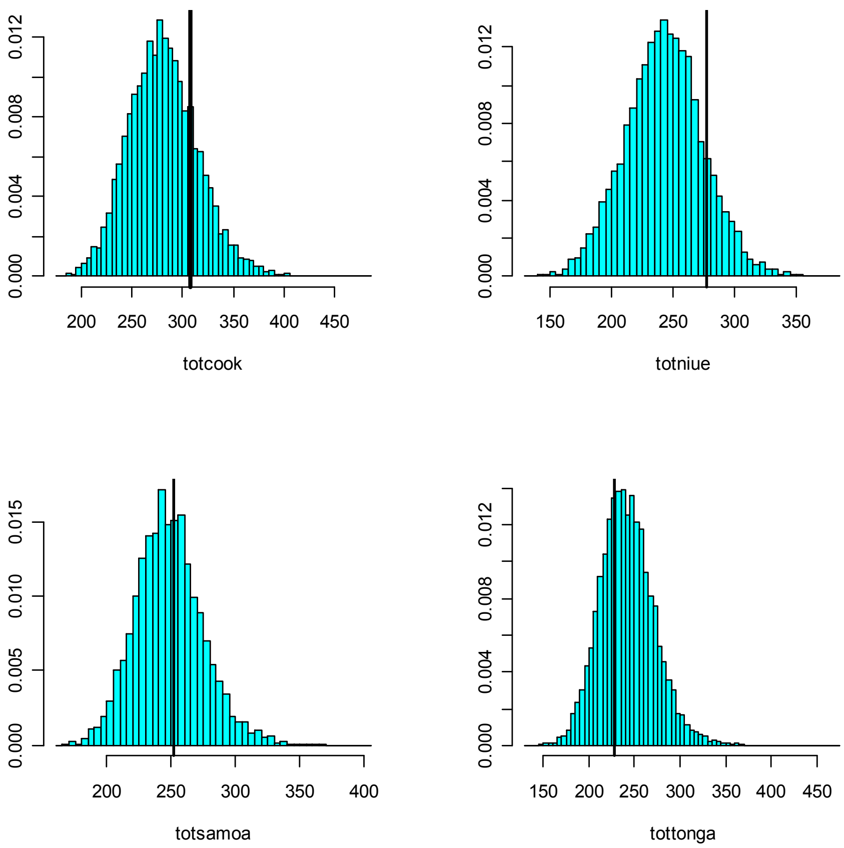
Cook Island mortality rates are greatest at 288 [95 percent CI 255 to 325] and Tongan rates smallest (234 [197, 269]), although there is substantial overlap in some CIs. However, the CI for the rate ratio (**Error! Reference source not found.**, Model 1) for Cook Island relative to Samoan mortality just excludes 1, providing modest support for the hypothesis that at least some of the differences are significant.

Figure 6
Posterior all-cause mortality rates for 0–74 year olds in 2001–04, sex and age adjusted



TC = total Cook Island Māori; TN = total Niuean; TS = total Samoan; TT = total Tongan.

Figure 7
Posterior predictive plots for age, sex standardised all-cause mortality rates from Model 1 (intercept varies by ethnicity)



Note: The vertical lines are empirical rates, standardised in the same way.

Figure 7 shows the observed age and sex standardised mortality rate (vertical line) compared to 7,000 simulations from the posterior predictive distribution for the model used in Figure 6. If the model provides a reasonable fit, then replicated data generated by the model should look similar to observed data. Hence for this model, there is no evidence for lack of fit, since empirical values for the observed rates are comfortably within the 95 percent range for the posterior predictive distribution. Adjusting additionally for and standardising by country of birth

A similar relative pattern is observed for posterior all-cause mortality rates from a model that included age, sex, and country-of-birth **Error! Reference source not found.** (Model 2) where, as before, standardisation used these same variables. The underlying data in this case were stratified by age, sex, country-of-birth, and Pacific ethnicity.

Posterior all-cause mortality rates are changed only slightly from Model 1, although overlap in CIs for Cook Island (295 [260, 333]) and Tongan (228 [193, 264]) all-cause mortality rates are reduced. Similar remarks can be made about Cook Island versus Samoan (243 [222, 264]) rates. However, the rate ratio (**Error! Reference source not found.**) for Cook Island relative to Samoan mortality excludes 1 more definitively, providing stronger support for the hypothesis that at least some of the differences are significant.

Adjusting additionally for income

There is very little change in the relative pattern for posterior all-cause mortality rates from Model 3 relative to Model 2. Similar comments can be made about rate magnitudes.

The overlap in credible intervals between Cook (294 [260, 332]) and Samoan (243 [222, 264]) rates is once again small and the rate ratio (**Error! Reference source not found.**, Model 3) for Cook Island relative to Samoan mortality excludes 1. For this model, the underlying data were stratified by age, sex, country-of-birth, income, and Pacific ethnicity.

Cause-specific mortality, 'Total' ethnicity

In this section, all (prior) models included age, sex, CoB, and income, and posterior rates were standardised using the same variables. As in Section **Error! Reference source not found.**, the underlying data were stratified by age, sex, CoB, income, and Pacific ethnicity. Posterior standardised rates for CVD mortality is shown in Figure 8 (below) and Table 13 (page 33). The relative pattern of mortality rates is also similar to that shown above, though absolute rates are necessarily smaller than the all-cause rates.

The Cook Island Māori CVD rate (111 [90, 135]) is 48 percent higher than the Niuean rate (75 [51,104]), 63 percent higher than the Samoan rate (68 [57, 79]), and 92 percent higher than the Tongan rate (58 [40, 78]). Furthermore, CIs for the latter two groups do not overlap with Cook Island Māori. This separation of CVD rates is also reflected in the rate ratio for Cook Island relative to Samoan (1.66 [1.26, 2.13]), shown in Table 13 and Figure 9.

Table 13

Posterior all-cause mortality rates (per 100,000) and rate ratios (95 percent credible intervals) by Pacific groups from models extended to include country of birth and household income using the restricted 2001–04 dataset

	Rates			Rate ratios c.f. Samoan		
Group	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
Cook Island	288 (255, 325)	295 (260, 333)	294 (260, 332)	1.18 (1.02,1.37)	1.22 (1.05, 1.41)	1.21 (1.05, 1.42)
Niue	252 (208, 297)	257 (211, 306)	260 (215, 308)	1.03 (0.84, 1.24)	1.06 (0.86, 1.28)	1.07 (0.88, 1.29)
Samoa	244 (224, 266)	243 (222, 264)	243 (222, 265)	1	1	1
Tonga	234 (197,269)	228 (193, 264)	226 (190, 262)	0.96 (0.80, 1.12)	0.94 (0.79, 1.11)	0.93 (0.77, 1.10)

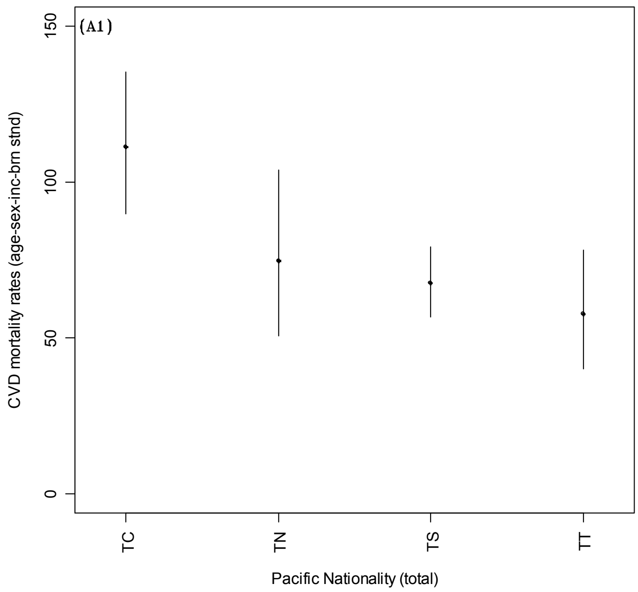
(1) Model 1 = data stratified by ethnic group, sex and age only; sex and age included as independent variables; posterior rates directly standardised to the NZ Pacific population using a total definition of ethnicity.
(2) Model 2 = Model 1, but extended to be additionally stratified by country of birth; additionally including country of birth as independent variable; posterior rates standardised as for Model 1.
(3) Model 3 = Model 2, but extended similarly for household income.

Table 13
Posterior all-cause and cause-specific mortality rates (per 100,000) and rate ratios (95 percent credible intervals) by Pacific groups, using the restricted 2001–04 dataset

	Rates per 100,000				Rate ratios c.f. Samoan			
Group	**All-cause**	**CVD**	**Cancer**	**Injury/ Suicide**	**All-cause**	**CVD**	**Cancer**	**Injury/ Suicide**
Cook Island	294 (260,332)	111 (90,135)	64 (50, 80)	31 (21, 42)	1.21 (1.05,1.42)	1.66 (1.26, 2.13)	0.85 (0.56, 1.09)	0.93 (0.60, 1.34)
Niue	260 (215,308)	75 (51,104)	66 (44, 88)	18 (6.0, 32)	1.07 (0.88,1.29)	1.11 (0.72,1.58)	0.87 (0.54, 1.19)	0.53 (0.17, 0.99)
Samoa	243 (222, 265)	68 (57, 79)	76 (65, 89)	33 (26, 42)	1	1	1	1
Tonga	226 (190,262)	58 (40, 78)	73 (56, 94)	27 (16, 41)	0.93 (0.77,1.10)	0.86 (0.58,1.20)	0.96 (0.71, 1.26)	0.82 (0.46, 1.28)

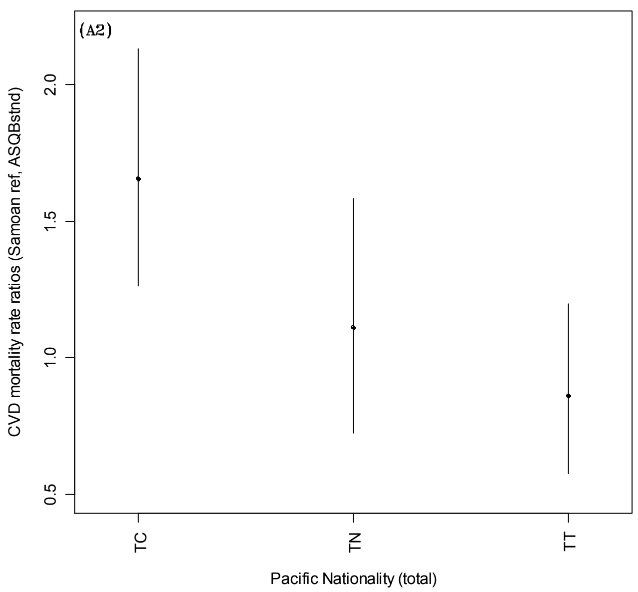
(1) HB models used data stratified by sex, age, country-of-birth, and income.
Standardisation was to the total Pacific population using a total definition of ethnicity.

Figure 8
Posterior CVD mortality rates for 0–74 year olds in 2001–04, sex, age, country of birth and income adjusted



TC = total Cook Island Māori; TN = total Niuean; TS = total Samoan; TT = total Tongan.

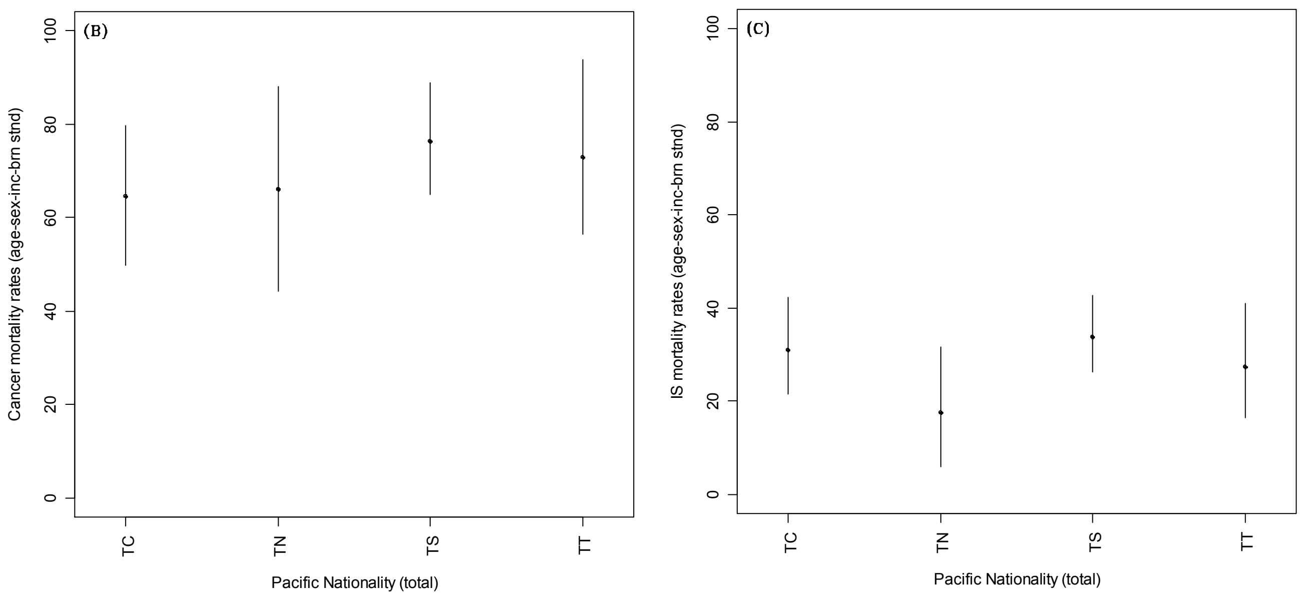
Figure 9
Posterior CVD mortality rate ratios for 0–74 year olds in 2001–04, sex, age, country of birth and income adjusted, compared to Samoan



TC = total Cook Island Māori; TN = total Niuean; TT = total Tongan.

Mortality rates were estimated for two other causes of death – cancer and injury/suicide (Figure 10). Mortality rates for cancer are similar across the four Pacific nationalities and CIs for the rate ratios all include unity. However, the Niuean injury/suicide mortality rate was nearly half of the Samoan rate with a 95 percent credible interval just excluding the null (rate ratio 0.53 (95 percent CI 0.17 to 0.99) – see Table 13.

Figure 10
Posterior cancer (B) and injury/suicide (C) mortality rates for 0–74 year olds in 2001–04, sex, age, country of birth and income



TC = total Cook Island Māori; TN = total Niuean; TS = total Samoan; TT = total Tongan.

Forcing a common prior intercept in HB modelling did not seem appropriate

In all the above analyses a separate intercept was allowed for each ethnic group in the prior model. We believe this was the best a priori theoretical position as we had good reason to believe mortality rates, especially CVD rates, would vary by Pacific group given census data (Table 3) and previous survey risk factor data – even if not well aligned with census data (**Error! Reference source not found.** and **Error! Reference source not found.**, (Sundborn et al 2008b)).

Nevertheless, given the lack of data on mortality differences between Pacific groups prior to this project, one might have conversely argued that there was not a strong a priori case for varying mortality rates. Had this been our prior position, we would have fitted models with a common intercept in the prior model for all four Pacific groups and allowed the observed data to 'pull' group-specific Pacific mortality rates away from the assumed common intercept if the empirical evidence was strong enough. Accordingly, we ran some HB models with a common intercept for CVD – the disease with the greatest likelihood of true variation.

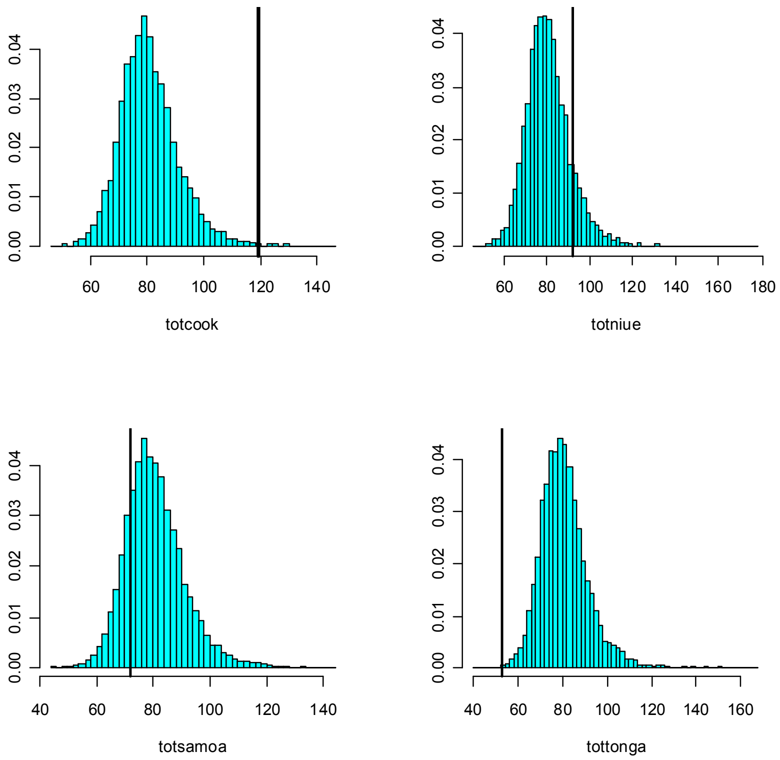
Figure 11 shows the distribution of posterior predicted mortality rates for the four Pacific groups compared to the empirically observed rate. Unlike the previous similar plot for HB models that allowed ethnic group-specific intercepts, empirical rates are less representative of the posterior predictive distribution, particularly for Cook and Tongan groups which suggests that the simpler common intercept prior is not a good assumption. This conclusion is strengthened by the deviance information criterion value for the fit (495.7) which is substantially larger than for the equivalent model where the intercept varies by ethnicity (450.6).

Similar results are found for all-cause mortality.

Taken together with the routinely standardised mortality rates in 1996–99 also showing higher mortality rates for Cook Island Māori, we believe allowing varying intercepts in the HB modelling is further justified. In the following discussion we focus attention on models where the prior intercept is allowed to vary across ethnic groups.

For the record, however, it is worth noting that the simpler (common intercept) model would still predict significant variations in CVD mortality by ethnicity, though less significantly than for the more complex model. This is, in of itself, an important sensitivity analysis because it says that even if the prior was that there were no differences in mortality between groups, the data cause us to re-consider. Thus, we have yet more confidence that the CVD rates are 'truly' different.

Figure 11
Posterior predictive plots for age, sex, CoB and income standardised mortality rates from Model 1 where the prior model intercept is forced to be the same across ethnic groups



Note: As before, the vertical line is the empirical rate.

All cause, 'Sole' ethnicity

Posterior standardised rates for 'Sole' ethnicity all-cause mortality rates from a prior model that included all predictors, and standardised for the same variables (Model 3) are shown in Table 14. As before, the underlying data were stratified by age, sex, country-of-birth, income, and Pacific ethnicity. Note that for consistency with the above HB results, standardisation is to the total rather than sole ethnicity population. As shown in Table 3, 66 percent, 53 percent, 71 percent and 41 percent of the total Samoan, Cook Island Māori, Tongan and Niuean populations were 'sole' (eg, Samoan only – no other Pacific ethnic group, no other ethnic group beyond Pacifika).

Table 14
Posterior 'sole' and 'total' all-cause mortality rates for 0–74 year olds in 2001–04, sex, age, CoB and income adjusted adjusted

	Rates		Rate ratios	
Group	**Sole: All-cause**	**Total: All-cause**	**Sole: All-cause**	**Total: All-cause**
Cook Island	302 (266, 344)	294 (260, 332)	1.18	1.21
Niue	296 (247, 360)	260 (215, 308)	1.16	1.07
Samoa	256 (231, 282)	243 (222, 265)	1.00	1.00
Tonga	249 (208, 289)	226 (190, 262)	0.97	0.93

Trends in posterior 'sole' mortality rates are similar to the 'total' ethnicity case, ie, rate ratios for the sole rate comparisons are much the same as for total rate comparisons (Table 14).

Sole Pacific posterior mean rates are larger than their total counterparts, though only slightly for Samoan and Cook Island ethnicities. There was considerable overlap of all CIs, and there is no evidence for significant differences between sole and total mortality rates at the level of individual ethnicities – although the pattern overall is to higher mortality rates among sole Pacific groups.

3.3 Sensitivity analyses

The results of basic sensitivity analyses about the sex and age adjusted (only) HB 0–74 year old mortality rates are shown in **Error! Reference source not found.**, and described in detail below. The sensitivity analyses are crude. For example, input parameters (eg, percentage returning to Pacific country) are applied to overall rates, not by strata of sex, age and so on.

What impact might differential return migration once ill have?

Return migration might occur when people are unwell, and they decide to return to their home country (country of birth, one would assume) to die. There is anecdotal and other evidence of this occurring in New Zealand. For example, survival from lung cancer among Pacific people appears good compared to other ethnic groups, yet this is due to a small amount of undercounting of Pacific deaths due to return migration (personal communication, Vladimir Stevanovic, Ministry of Health, 2007).

In parallel with this project, Tobias and others of Health and Disability Intelligence at the Ministry of Health attempted to estimate how much Pacific mortality rates might be underestimated due to return migration (personal communication, Martin Tobias, Ministry of Health, 2008). Using lung cancer as a 'tracer condition', they estimated that Pacific mortality rates in the early 2000s might be underestimated by about 8 percent. These calculations involved determining the lung cancer mortality from New Zealand Health Information Service data, noting it was less than expected, estimating what that lung cancer mortality should have been, assuming that other chronic diseases were similarly underestimated due to return migration, and then calculating adjustment factors for overall Pacific mortality having dichotomised deaths into those subject to return migration (ie, chronic conditions) and those not (ie, acute conditions).

We set 8 percent as the best estimate of underestimation of Pacific deaths due to return migration, and 4 percent and 12 percent as low and high scenarios. When such bias is the same for all four Pacific ethnic groups, the rates vary but rate ratios do not (**Error! Reference source not found.**), meaning no bias in the rate ratios estimated in this project.

However, would return migration when terminally ill be the same for the four Pacific groups in this project? We thought that return migration would be more likely among Tongan and Samoan people because: there are still substantial Samoan and Tongan populations (and health care facilities) in the home Island to act as 'pull factors' for return migration (eg, see Table 1); and lower proportions of Samoan and Tongan people are born in New Zealand, probably predicting greater return migration for these groups when terminally ill. Scenario A (10 percent of terminally ill Samoan and Tongan people returning home to die, and 5 percent of Niuean and Cook Island Māori returning home to die when terminally ill) is our best estimate of such differential bias. For all cause-mortality, this would reduce the Cook Island: Samoan excess rate ratio by about a quarter from 1.21 to 1.15, but would only reduce the CVD rate ratio by about 13 percent from 1.63 to 1.55.

The two remaining differential return migration scenarios (B = 12 percent Samoan and 3 percent Cook Island Māori; C = 25 percent Samoan; 0 percent Cook Island Māori) are both in our view unlikely if not implausible. However, their purpose is to determine what amount of differential return migration bias would be needed to cause the observed higher mortality for Cook Island Māori. That is, we are determining how much bias would be required to render the findings null. Given that, in our view, the amount of bias required to 'nullify' the key findings is implausible, we conclude differential return migrant bias is an implausible cause of differences in CVD mortality.

Table 15
Sensitivity analyses about HB all-cause and CVD mortality rates in case of; return migration when terminally ill; census undercounting

		Rates				Rate ratios			
		Cook	Niuean	Samoan	Tongan	Cook	Niuean	Samoan	Tongan
HB all-cause rate and rate ratios		294	260	243	226	1.21	1.07	1	0.93
Amount of return migration for terminally ill									
Non-differential	Best estimate = 8%	320	283	264	246	1.21	1.07	1	0.93
	Low scenario = 4%	306	271	253	235	1.21	1.07	1	0.93
	High scenario 12%	334	295	276	257	1.21	1.07	1	0.93
Differential, about best estimate of 8% as average	A. 10% Samoan and Tongan; 5% Niuean and Cook	309	274	270	251	1.15	1.01	1	0.93
	B. 12% Samoan; 3% Cook	303		276		1.10		1	
	C. 25% Samoan; 0% Cook	294		324		0.91		1	
Amount of census undercount									
Non-differential	Best estimate = 4%	282	250	233	217	1.21	1.07	1	0.93
	Low scenario = 2%	288	255	238	221	1.21	1.07	1	0.93
	High scenario = 6%	276	244	228	212	1.21	1.07	1	0.93
Differential, about best estimate of 4% as average	6% Samoan and Tongan; 2% Niuean and Cook	288	255	228	212	1.26	1.12	1	0.93
HB CVD rate and rate ratios		111	75	68	58	1.63	1.10	1	0.85
Amount of return migration for terminally ill									
Non-differential	Best estimate = 8%	121	82	74	63	1.63	1.10	1	0.85
	Low scenario = 4%	116	78	71	60	1.63	1.10	1	0.85
	High scenario 12%	126	85	77	66	1.63	1.10	1	0.85
Differential, about best estimate of 8% as average	A. 10% Samoan and Tongan; 5% Niuean and Cook	117	79	76	64	1.55	1.04	1	0.85
	B. 12% Samoan; 3% Cook	114		77		1.48		1	
	C. 25% Samoan; 0% Cook	111		91		1.22		1	
Amount of census undercount									
Non-differential	Best estimate = 4%	107	72	65	56	1.63	1.10	1	0.85
	Low scenario = 2%	109	74	67	57	1.63	1.10	1	0.85
	High scenario = 6%	104	71	64	55	1.63	1.10	1	0.85
Differential, about best estimate of 4% as average	6% Samoan and Tongan; 2% Niuean and Cook	109	74	64	55	1.70	1.15	1	0.85

What impact might differential census under-counting have?

Not everyone who should is enumerated by the New Zealand census. Issues surrounding ‘Pacific over-stayers’ and ‘dawn raids’ by immigration authorities in the 1970s have probably made Pacific people less inclined to cooperate with census enumeration.

The 2001 post enumeration survey (PES) found that the 2001 census underestimated the total eligible New Zealand resident population by about 2.2 percent, but underestimated the Pacific population by 5.2 percent (Statistics New Zealand 2002). For the purposes of this project, we need an estimate of the Pacific undercount among the older Pacific population where the majority of deaths occur. As much of census undercounting is driven by younger people (and the Pacific population is particularly young), it seems reasonable to expect that the ‘weighted average’ census undercount for the Pacific population at risk of dying is less than 5.2 percent. On the other hand, the consensus of the authors of this report is that the PES probably underestimates the true undercount – the PES and census are unlikely to be independent sampling frames. For example, the same field staff are used in the PES and the capture-recapture effect is well known as a limitation in this type of exercise.

What will census under-enumeration mean for calculation of mortality rates in the NZCMS? It will mean that some of the Pacific deaths we have on the mortality file that we fail to link to the census file will not be due to ‘failure’ of the linkage but actually due to not being present in the census file. The linkage bias adjusters in the NZCMS assume that these deaths truly had their corresponding census record present in the census file, and calculate linkage bias weights to apply to all linked mortality records to make them representative of all eligible mortality records.(Fawcett et al 2008) If there is non-trivial under-enumeration of Pacific people (especially among older age groups), then NZCMS Pacific mortality rates will overestimate Pacific mortality.

The next point to note is that the relevant percentage undercount is not for all census respondents, but one weighted to the age distribution of deaths (census under-enumeration varies by age). Using the census distribution of Pacific people, and of deaths, and applying percentage undercounts of age that fit the overall pattern by age in the 2001 census PES yet give the Pacific census undercount of 5.2 percent, the percentage undercount for a likely age distribution of deaths was 3.1 percent (calculations available from authors on request).

For the purposes of this report we thought this 3.1 percent was still probably low due to deficiencies in the PES study design. Accordingly we selected a 4 percent undercount as an overall estimate, and 2 percent and 6 percent as low and high scenarios about this estimated census undercount. With respect to the calculation of mortality rates, these percentages cause an *overestimate* of the mortality rate. Pacific mortality rates adjusted for such overestimation, when non-differential across the four Pacific groups, have no impact on the rate ratios compared to Samoan (**Error! Reference source not found.**) – just the rates themselves.

Is it plausible that census undercounting might vary between the four Pacific groups, and hence overestimation of mortality rates might vary by Pacific group? As Niuean and Cook Island Māori have automatic New Zealand citizenship, there might be less willingness among Samoan and Tongan population to be enumerated due to the history of action against ‘over-stayers’. We posited a 2 percent census undercount for Niuean and Cook Island Māori, and 6 percent for Samoan and Tongan people. However, this only serves to widen the mortality gap between Cook Island Māori and Samoan people (**Error! Reference source not found.**). Therefore, we conclude that differential census under enumeration across the four Pacific groups in our study is very unlikely (if not impossible) to cause the observed discrepancies in mortality rates.

What impact might terminally ill Pacific People migrating to New Zealand to treatment have?

It is anecdotally noted that non-New Zealand resident Pacific people may come to New Zealand when terminally ill, then die in New Zealand, 'spuriously' increasing the apparent Pacific mortality rate in New Zealand. There are three relevant variants on this scenario to consider for NZCMS analyses.

First, consider a terminally ill Pacific person who migrates to New Zealand after 2001 census night, and dies before March 2004. As they arrived after census night, they will not be in the census file. If they are not a New Zealand resident at time of death, we exclude them from the eligible mortality files. If they were recorded as New Zealand resident (correctly or incorrectly), they should still be excluded from the eligible mortality files using the 'years in New Zealand' variable on the mortality file. (This latter assumption depends on the accuracy of this variable.)

Second, consider a terminally ill Pacific person who migrates to New Zealand before 2001 census night, and dies before March 2004. As long as they were not recorded as a New Zealand resident on the census or mortality file, no bias should result.

Third, consider a terminally ill Pacific person who migrates to New Zealand before 2001 census night, is a New Zealand resident on census night 2001, and dies before March 2004. They should be counted in both the census and mortality file. But it is disputable whether this is a bias as they are eligible due to New Zealand residency. If this scenario varies notably between Samoan, Niuean, Tongan and Cook Island Māori, variations in mortality rate may arise. For example, one might argue that because all Niuean and Cook Island Māori people are New Zealand citizens, greater flow in and out of people during illness may occur, possibly increasing the Cook Island Māori and Niuean mortality rate recorded in the NZCMS.

We have not explicitly conducted sensitivity analyses about these above scenarios and variants. There is a paucity of information, be it accuracy of data on residency on census and mortality data, or frequent/recent migration by terminally ill Pacific people who are also New Zealand residents. However, we believe that the biases are not likely to be large. For example, under the third variant above that will spuriously increase Niuean and Cook Island Māori mortality rates, it must be noted that 14 and five times as many people of these ethnic groups (respectively) live in New Zealand than do in the home country (Table 1, page 12). Therefore, mortality rates will be numerically driven by those living in New Zealand – not those migrating to New Zealand for disease reasons from smaller home country populations.

What impact might differential error in NZCMS linkage weights have?

As mentioned above, the NZCMS fails to link all eligible mortality records back to a census record. A main reason is residential mobility between census night and death. All routine analyses of NZCMS data, therefore, are weighted analyses whereby linked mortality records are weighted up (and unlinked census respondents slightly weighted down). In generating these weights, we included strata of Māori, Pacific, Asian and nMnPnA. However, we did not specifically assess the possibility of differential linkage success between Pacific groups. It was beyond the scope of this study, and arguably beyond the scope of data stability, to conduct linkage weights separately by Pacific group. Thus, we cannot rule out the possibility of some residual linkage bias between the four Pacific groups in this study. Nevertheless, we do not believe if present it would be large – and certainly not large enough to account for the higher Cook Island Māori mortality rates.

3.4 Summary of sensitivity analyses

Return migration of Pacific people to their home country upon learning of a terminal diagnosis might be differential by Pacific group, and might explain some of the apparent excess all-cause mortality for Cook Island Māori compared. However, even under extreme assumptions it would account for only a negligible amount of the excess CVD mortality of Cook Island Māori.

Other possible differential biases across the four Pacific groups, such as census under enumeration, migration to New Zealand with terminal illness, and residual linkage bias in the NZCMS, seem almost certainly not able to generate the observed higher mortality for Cook Island Māori.

As shown previously in this report, the differences in mortality rates between Niuean, Samoan, and Tongan people are smaller in magnitude, and rankings in mortality change for varying data restrictions and census cohorts (see Table 10). Furthermore confidence intervals and credible intervals substantially overlap – although one exception might be an apparently lower injury/suicide mortality rate for Niuean people. Thus, it is plausible that small amounts of the above systematic biases, especially if having a cumulative or interacting effect, might alter rankings of mortality rates among these three Pacific groups. But, again, the wide confidence and credible intervals reflecting random errors and non-quantifiable systematic uncertainties in the data are probably more important impediments to accuracy and precision than residual systematic biases considered here.

4 Discussion

4.1 Key findings

Cook Island Māori had elevated all-cause and CVD mortality compared to the three other Pacific groups.

For all-cause mortality, and using the HB estimates for the sex and age adjusted model (‘total’ ethnic groups, sexes pooled, 2001–04 cohort, 0-74 year olds), the Cook Island Māori mortality rate was 1.18 (95 percent credible interval 1.02 to 1.37) times greater than the Samoan rate. Equivalent rate ratios for Niuean and Tongan compared to Samoan were 1.03 (0.84 to 1.24) and 0.96 (0.80 to 1.12). Adjusting additionally for country of birth and equivalised household income did not make a material difference to these estimates: the Cook Island rate ratio was 1.21 (1.05 to 1.42); Niuean 1.07 (0.88 to 1.29); Tongan 0.93 (0.77 to 1.10).

For CVD mortality, and using the HB estimates for the sex, age, country of birth and income adjusted model, the Cook Island Māori mortality rate was 1.66 (1.26 to 2.13) times greater than the Samoan rate. Equivalent rate ratios for Niuean and Tongan compared to Samoan were 1.11 (0.72 to 1.58) and 0.86 (0.58 to 1.20).

Analyses for ‘sole’ ethnic groups using HB modelling, and routine direct standardisation analyses on both the 2001-04 and 1996-99 cohorts, were consistent with Cook Island Māori having moderately elevated all-cause mortality, and substantially elevated CVD mortality, compared to the other more equivalent Pacific groups.

Furthermore, a variety of sensitivity analyses about the likely residual systematic biases we could posit (return to country of origin migration when terminally ill; census undercounting; residual NZCMS linkage bias; migration from Pacific Islands to New Zealand when sick) suggest that the elevated mortality rate for Cook Island Māori (especially for CVD) was robust.

Under HB modelling, cancer mortality did not vary between the four Pacific groups either substantially or significantly but injury/suicide mortality among Niuean was 0.53 (0.17 to 0.99) of that among Samoans with no substantive variation among the three other ethnic groups.

To our knowledge, this project is the first time in New Zealand that clear (and marked in the case of CVD) differences in mortality have been demonstrated between Pacific people. There is parallel emerging evidence from the Ministry of Health on difference in youth health status (work in progress) and some indicators of adult health status (Health and Disability Intelligence, Ministry of Health, work in progress), and published evidence of diversity for mental health.(Ministry of Health 2006). This diversity between Pacific populations is important for further health research and policy. Researchers should, if at all possible, try and analyse and present separate results for appropriate Pacific groupings. Policymakers increasingly need to consider Pacific group-specific policies and programmes, especially for CVD.

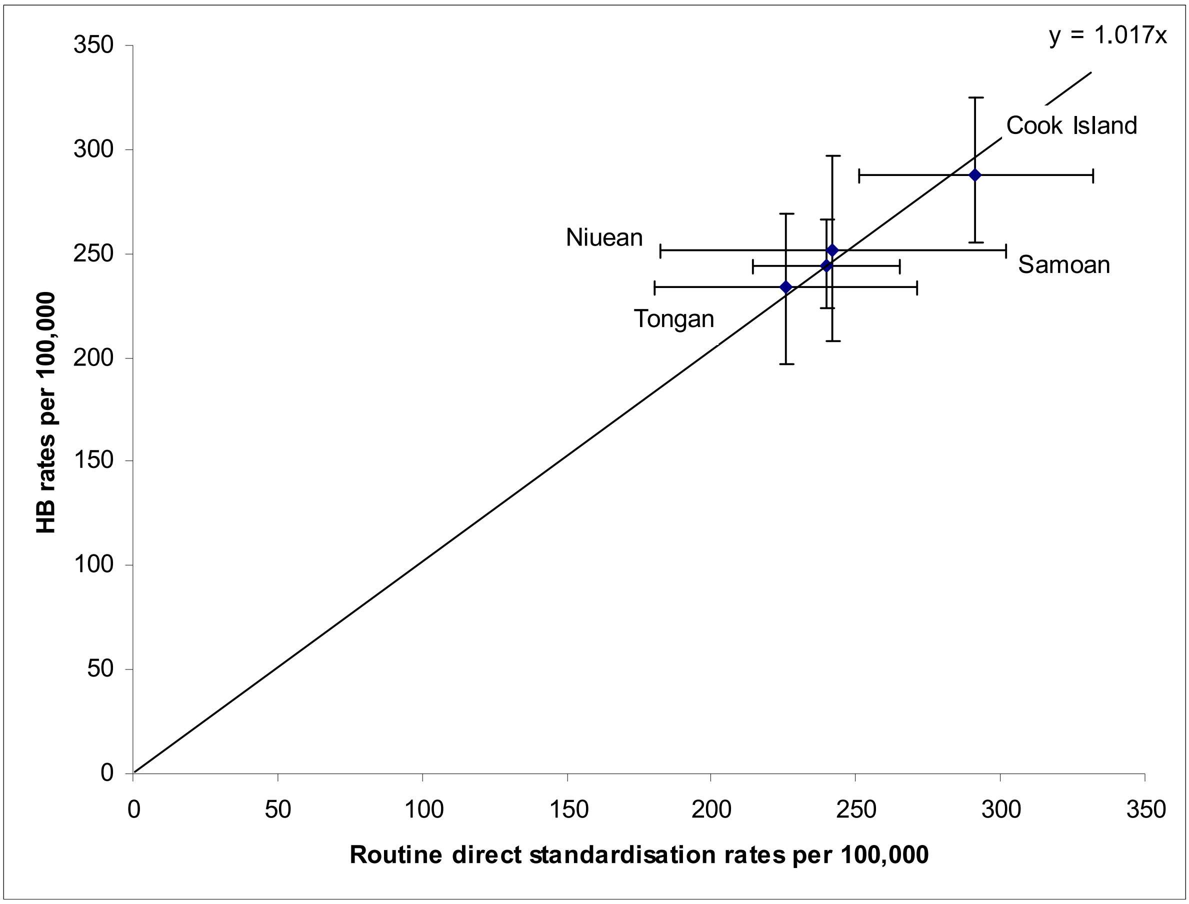
4.2 Hierarchical Bayesian modelling: lessons from this project for future applications

We successfully implemented HB modelling on linked census-mortality data to assess Pacific group-specific mortality rate variations. Significant researcher up-skilling was required, detailed cross-classified data had to be extracted from the Data Laboratory, and then HB modelling had to be undertaken off-site from the Data Laboratory. For future projects, some of these barriers will be reduced – especially the research-team skills. It is worthwhile considering the advantages of this approach to this (and other) ‘small-cell’ problems. But first, we directly compare the HB modelling output and that from ‘routine’ direct standardisation.

How did ‘routine’ direct standardisation and HB modelling results compare?

In this project the application of HB modelling made little notable difference over the more straightforward application of routine direct standardisation methods. Figure 12 shows a plot of the HB versus routine direct standardisation estimates for all-cause mortality among the restricted 0–74 year old cohort. The regression line is for a simple unweighted OLS regression with a forced zero y-axis intercept. The slope of the regression line is 1.017, indicating that the HB estimates tended to be slightly higher. But what really stands out is that the estimates largely fall very close to the regression line. Perhaps the only noteworthy shift is for Cook Island Māori, where the HB estimate is less than the routine direct standardisation estimate. This is logical: as expected HB modelling has shrunk the particularly high Cook Island Māori estimate slightly towards the three other Pacific groups.

Figure 12
Scatter plot of HB versus direct standardisation all-cause mortality rates, sexes combined, both calculated as similarly as possible on restricted 2001–04 datasets for 0–74 year olds adjusted for sex and age



One apparent advantage in the above plot for HB estimates over the routine direct standardisation estimates was narrower credible intervals. However, we did not specifically examine possible reasons for such differences between credible and confidence intervals. It must be noted that the NZCMS routine direct standardisation procedures also routinely include some variance inflation to allow for application of linkage bias weights – the HB modelling was not extended to do the same. Thus, extra analytical work may find that the HB credible intervals need to be inflated somewhat. On the other hand, simplifying assumptions are required to derive errors and confidence intervals in the direct method and biased estimates of uncertainty are common in such situations (Allison 2005; Greenland 2007). This is particularly true for over-dispersed count data (Allison 2005; Greenland 2007), as is likely to be the case for the data used here. It might be a useful future line of work to validate and compare each of the confidence intervals for routine direct standardisation and credible intervals from an HB approach.

Overall, the HB modelling in this project gives us greater confidence that there truly are differences in mortality rates between Pacific groups. However, it would appear that the reporting of mortality rates for all ages combined for the groupings we choose *are not yet sparse enough data for the HB modelling estimates to differ from routine standardisation*. Put another way, this particular project did not lie in the 'zone' where routine analyses become too unstable due to sparse data, yet HB analyses retain good predictive accuracy. Rather, we have validated the application of HB modelling.

Future possible applications of HB modelling

Theoretically, HB modelling has clear advantages. It allows an a priori model to be posited, yet allows posterior estimates to deviate from such parametrically predicted models if the observed data substantially varies from the predicted model and there are sufficient observations.

The results presented in this report are at a high level of aggregation, and (in retrospect) we would have been concerned had there been gross differences between empirical and HB rates (as in Figure 11). The strength of the method should be more apparent if we wish to examine less aggregated results; eg, rate ratios over age-sex-CoB strata. Young et al took this approach with NZCMS data and demonstrated an improvement in estimates of mortality rates in data-poor strata for ethnicity by education relative to standard approaches (Young et al 2006). It is interesting to note, for example, that in Table 3 mortality rates for Cook Island Māori and Niuean ethnicities increase slightly when CoB is included in the analysis (Model 2 versus Model 1) as might be expected given the higher proportion born in New Zealand. However, after standardising posterior rates, the difference is small and a direct comparison of rates across CoB strata might be more informative.

It is worth noting that the direct approach, which estimates empirical mean rates in each stratum, are closely related to saturated (1 parameter per stratum) models. For the particular case of an age-sex model, a saturated model requires 40 parameters and maximum likelihood estimates are the empirical rates. This model achieves a good fit to a given dataset, but usually at the expense of out-of-sample predictive ability, essentially because it over-fits to the random component. In contrast, the HB age-sex model requires about 24 parameters and we would expect significantly better out-of-sample predictive behaviour (eg, predicting mortality rates for older Pacific people, or Pacific people in other time periods).

Regarding future applications of HB modelling, we note three domains of potentially productive future work:

1. Extending HB modelling on *NZCMS data for Pacific people*:
    a. to examine/predict variations between these four Pacific groups by age group, sex, country of birth, and income-level (ie, disaggregating the overall estimates provided in this report)
    b. to include smaller Pacific groups (eg, Tokelauan)
    c. to use other earlier NZCMS cohorts as priors to improve temporal estimations in 2001–04, and predictions thereafter
    d. to include the smoking variable in the 1996–99 cohort to assist with predicting future pacific mortality burden as smoking prevalence changes
    e. to explore possible variations in mortality between people identifying as just one Pacific group, two or more Pacific groups or other multiple ethnic combinations
    f. to explore the influence of immigration status and acculturation processes on Pacific subgroup mortality rates.
2. Extending HB modelling on *NZCMS and other health data*:

   a. There is concern that the obesity epidemic might see an upturn in ischaemic heart disease mortality among younger cohorts, and indeed some preliminary evidence of this in New Zealand (Tobias et al 2006). HB modelling might be useful in exploring this (currently) sparse data problem, by age and socioeconomic position.
   b. More people are self-identifying as two or more ethnic groups, and there appear to be pronounced differences in mortality between people of sole and multiple groups (Callister and Blakely 2004). HB modelling might be a useful method to further understand health variations by complex ethnic groups, and possibly for emerging groups (eg, New Zealander).

3. *Official statistics.* Bayesian approaches have already been applied to imputation, complex survey design, and the creation of synthetic datasets. Further applications of HB modelling to official statistics might include:
   a. Surveys such as the Household Labour Force Survey, Household Economic Survey, and Time Use Survey which have small numbers of people in particular ethnic groups
   b. Causal inference with repeated measures data such as the Survey of Family, Income and Employment.

For the ‘small domain’ regression problem, such as the one considered here, hierarchical Bayesian methods can have an important role in the toolkit of quantitative analysis. Shrinkage estimators often improve estimation and prediction accuracy (Greenland 2007). Gelman (Gelman 2006) argues that multi-level models are almost always an improvement over classical regression, though to varying degrees depending on whether one is interested in prediction (can be essential), data reduction (can be useful), or causal inference (can be helpful). Analysis of longitudinal data and of missing and misclassified data using conditional independence models is another area of significant application for Bayesian approaches (Ashby 2006).

It must also be noted, especially by managers and funders, that there is an entry cost to using HB modelling, the need to train staff, and the need to be able to distil interpretations back to messages understandable to the target audience. But HB modelling is only likely to increase in usage in the future as the emphasis on making interpretations in the context of sparse data increases, and the method becomes more ‘mainstream’. Quantitative analysts and statisticians within and without Statistics New Zealand will need to increase their capacity and competence in HB methods.

References

Allison, PD (2005). *Fixed Effects Regression Methods for Longitudinal Data: Using SAS*, SAS Publishing.

Anae, M, Luli, L, & Burgoyne, L (2006). *The Polynesian Panthers 1971–1974: the crucible years*, Reed, Auckland.

Ashby, D (2006). 'Bayesian statistics in medicine: a 25 year review', *Statistics in Medicine*, 25, 3589–3631.

Bedford, R (2008). 'Pasifika Mobility: Pathways, circuits, and challenges in the 21st century', in *Pacific Interactions: Pasifika in New Zealand, New Zealand in Pasifika*, ed A. Bisley, Institute of Policy Studies, Wellington.

Best, N, Richardson, S, & Thomson, A (2005). 'A comparison of Bayesian spatial models for disease mapping', *Statistical Methods in Medical Research*, 14, 35.

Blakely, T, Tobias, M, Atkinson, J, Yeh, L-C, & Huang, K (2007). 'Tracking Disparity: Trends in ethnic and socioeconomic inequalities in mortality, 1981–2004', Ministry of Health, Wellington.

Brooks, SP & Gelman, A (1998). 'General methods for monitoring convergence of iterative simulations', *Journal of Computational and Graphical Statistics*, 7, 434–455.

Callister, P & Blakely, T (2004). 'Ethnic classification, intermarriage, and mortality: Some methodological issues in relation to ethnic comparisons in Aotearoa/New Zealand' [Working Paper], Wellington School of Medicine and Health Sciences, University of Otago, Wellington.

Callister, P & Didham, R (2008). 'Emerging demographic and socioeconomic features of the Pacific population in New Zealand', in *Pacific Interactions: Pasifika in New Zealand, New Zealand in Pasifika*, ed A. Bisley, Institute of Policy Studies, Wellington.

Christiansen, CL, & Morris, CN (1997). 'Hierarchical Poisson regression modeling', *Journal of the American Statistical Association*, 92, 618–632.

Cook, L, Didham, R, & Khawaja, M (1999). 'On the demography of Pacific people in New Zealand', Statistics New Zealand, Wellington.

Daniels, MJ (1999). 'A Prior for the Variance in Hierarchical Models', *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 27, 567–578.

Fairburn-Dunlop, P, & Makisi, GS (2003). *Making our place: growing up PI in New Zealand*, Dunmore Press, Palmerston North.

Fawcett, J, Atkinson, J, & Blakely, T (2008). 'Record linkage of census and mortality 2001–04 records', *New Zealand Census-Mortality Study Technical Report No. 6*. Health Inequalities Research Programme, University of Otago, Wellington.

Gelman, A (2006). 'Multilevel (hierarchical) modeling: what it can and can't do', *Technometrics*, 48, 432–435.

Gelman, A, Carlin, JB, Stern, HS, & Rubin, DB (2004). *Bayesian data analysis*. Chapman & Hall/CRC, Boca Raton, Florida.

Greenland, S (2007). 'Bayesian perspectives for epidemiological research. II. Regression analysis', *Int. J. Epidemiol*, 36, 195–202.

— (2008). 'Invited Commentary: Variable Selection versus Shrinkage in the Control of Multiple Confounders', *American Journal of Epidemiology*, 167, 523.

King, M (2003). *The Penguin History of New Zealand*, Penguin, Auckland.

Macpherson, C (2004). 'From Pacific Islanders to Pacific People: The Past, Present and Future of the Pacific Population in Aotearoa', in *Tangata, Tangata: The Changing Contours* *of Ethnicity in Aotearoa New Zealand*, eds P Spoonley, C Macpherson & D Pearson, Thomson Dunmore Press, Wellington.

— (2006). 'Pacific People in Aotearoa/New Zealand: from Sojourn to Settlement', *Migration Happens. Reasons, Effects and Opportunities of Migration in the South Pacific*, in eds K Ferro & M Wallner, Lit Verlag GmbH & Co. KG, Vienna.

Ministry of Health (2008). 'A Portrait of Health: Key Results of the 2006/07 New Zealand Health Survey', Ministry of Health, Wellington.

Ministry of Health & Ministry of Pacific Island Affairs (2004). 'Tupu ola moui: Pacific health chartbook', Ministry of Health, Wellington.

Spiegelhalter, DJ, Best, NG, Carlin, BP, & Van Der Linde, A (2002). 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society, Series B, statistical methodology*, 64, 583–639.

Statistics New Zealand (2002). 'A Report on the Post-enumeration Survey 2001', Statistics New Zealand, Wellington.

Sundborn, G, Metcalf, P, Gentles, D, Scragg, R, Schaaf, D, Dyall, L, Black, P, & Jackson, R (2008a). 'Ethnic differences in cardiovascular disease risk factors and diabetes status for Pacific ethnic groups, and Europeans in the Diabetes, Heart and Health Survey (DHAH), 2002–2003', *New Zealand Medical Journal*, 121, 28–39.

— (2008b). 'Ethnic differences in cardiovascular disease risk factors and diabetes status for Pacific ethnic groups, and Europeans in the Diabetes, Heart and Health Survey (DHAH), 2002–2003', *New Zealand Medical Journal*, 121, 28–39.

Tobias, M, Sexton, K, Mann, S, & Sharpe, N (2006). 'How low can it go? Projecting ischaemic heart disease mortality in New Zealand to 2015' [see comment], *New Zealand Medical Journal*, 119, U1932.

Young, J, Graham, P, & Blakely, T (2006). 'Modeling the relation between socioeconomic status and mortality in a mixture of majority and minority ethnic groups', *American Journal of Epidemiology*, 164, 282–291.

Appendix

Table 16
Comparison of CVD risk factors among *men* aged 35-74 years (age adjusted). Values are mean (SE) or percent

Variables	European (n=863)	Samoan (n=246)	Tongan (n=123)	Niuean (n=49)	Cook Island (n=46)
Age in years	51.7 (0.18)	46.6 (0.77)***	53.8 (2.51) †	48.0 (1.52)*	50.0 (4.26)
Current Smokers (%)	15.2% (0.01)	34.2% (0.05)***	36.3% (0.08)**	18.4% (0.07)	19.2% (0.06)
Never Smokers (%)	46.8% (0.02)	30.4% (0.05)***	25.5% (0.08)*	41.6% (0.08)	49.3% (0.12)
Inactive leisure time (%)	24.3% (0.02)	45.2% (0.05)***	39.3% (0.08)	29.1% (0.07)	47.2% (0.10)*
Exercise (min/week) (tolerance)	66 (1.18)	26 (1.91)**	29 (2.67)	56 (2.22)	33 (2.36)
NZDep2001	4.4 (0.10)	7.9 (0.24)***	8.7 (0.30)***	7.9 (0.34)***	7.6 (0.45)***
BMI (kg/m²)	27.6 (0.16)	33.0 (0.47)***	34.4 (0.88)***	32.0 (0.75)***	30.8 (0.81)*** †
Height (cm)	176.7 (0.25)	173.5 (0.69)***	174.7 (0.85)*	173.0 (1.04)***	172.3 (1.78)*
Waist-hip ratio	0.93 (0.002)	0.94 (0.01)	0.97 (0.01)*** †	0.93 (0.01)	0.93 (1.01)
Waist (cm)	97.8 (0.39)	105.9 (1.11)***	110.0 (2.05)*** †	102.1 (1.57)** †	100.2 (2.93)
Blood pressure (mmHg)					
Systolic	126 (0.62)	133 (2.44)**	125 (2.57)	129 (2.55)	135 (3.24)**
Diastolic	78 (0.35)	83 (1.26)***	81 (1.55)	79 (1.67)	81 (1.81)
Serum lipids (mmol/L)					
Total cholesterol	5.61 (0.04)	5.28 (0.08)***	5.62 (0.19)	5.36 (0.17)	5.72 (0.23)
HDL cholesterol	1.31 (0.01)	1.24 (0.03)*	1.12 (0.04)*** †	1.31 (0.05)	1.27 (0.05)
Ratio Tot/HDL	4.51 (0.05)	4.52 (0.14)	5.18 (0.24)** †	4.26 (0.17)	4.63 (0.17)
Triglycerides	1.65 (0.04)	1.68 (0.12)	1.67 (0.19)	1.51 (0.16)	1.56 (0.14)
LDL-cholesterol	3.57 (0.03)	3.31 (0.09)**	3.77 (0.17) †	3.38 (0.15)	3.74 (0.22)
Microalbuminuria (mg/L)	4.11 (1.06)	7.30 (1.22)***	4.45 (1.28) ††	5.53 (1.46)	7.26 (1.78)
Five-year risk score of CVD	6.8% (0.15)	9.4% (0.70)***	10.8% (1.13)***	7.1% (0.37) ††	9.1% (0.54)***

*0.01<p<0.05, **0.001<p<0.01, ***p<0.001 compared to European; †0.01<p<0.05, ††0.001<p<0.01,
†††p<0.001 compared to Samoan, = geometric mean
Source: Sundborn et al, 2008

Table 17
Comparison of CVD risk factors among *women* aged 35-74 years (age adjusted). Values are mean (SE) or percent

Variables	European (n=863)	Samoan (n=246)	Tongan (n=123)	Niuean (n=49)	Cook Island Maori (n=46)
Age in years	52.1 (0.18)	46.9 (0.80)***	48.0 (1.67)*	46.2 (1.20)***	46.1 (1.02)***
Current Smokers (%)	13.3% (0.01)	18.0% (0.04)	7.5% (0.03) †	12.7% (0.06)	30.5% (0.09)
Never Smokers (%)	53.9% (0.02)	54.8% (0.06)	84.0% (0.05)*** †††	49.6% (0.11)	38.4% (0.08)*
Inactive leisure time (%)	23.3% (0.02)	37.7% (0.05)**	47.0% (0.07)**	36.6% (0.09)	45.8% (0.08)*
Exercise (min/week) (tolerance)	59 (1.16)	23 (1.57)***	15 (2.31)**	37 (2.33)	17 (2.41)*
NZDep2001	4.5 (0.11)	8.5 (0.27)***	9.0 (0.16)***	8.0 (0.38)***	7.9 (0.63)***
BMI (kg/m²)	27.3 (0.19)	36.3 (0.57)***	36.3 (1.05)***	35.3 (1.27)***	35.4 (0.78)***
Height (cm)	163.5 (0.23)	160.8 (0.51)***	165.1 (0.87) †††	162.8 (0.76) †	160.3 (1.25)*
Waist-hip ratio	0.81 (0.002)	0.87 (0.01)***	0.87 (0.01)***	0.85 (0.01)*** ††	0.87 (0.01)***
Waist (cm)	87.1 (0.45)	105.1 (1.06)***	106.4 (1.73)***	101.0 (1.80)***	102.7 (2.00)***
Blood pressure (mmHg):					
Systolic	116 (0.60)	128 (1.55)***	121 (3.85)	126 (2.64)***	128 (2.53)***
Diastolic	72 (0.34)	78 (1.00)***	78 (1.59)***	79 (1.57)***	82 (2.97)**
Serum lipids (mmol/L):					
Total cholesterol	5.50 (0.03)	5.40 (0.12)	5.45 (0.13)	5.01 (0.09)*** †	5.88 (0.34)
HDL cholesterol	1.62 (0.01)	1.36 (0.03)***	1.31 (0.05)***	1.39 (0.06)***	1.31 (0.04)***
Ratio Tot/HDL	3.57 (0.04)	4.07 (0.10)***	4.36 (0.21)***	3.77 (0.15)	4.72 (0.40)**
Triglycerides	1.20 (0.02)	1.30 (0.06)	1.39 (0.08)*	1.34 (0.08)	1.78 (0.27)*
LDL-cholesterol	3.33 (0.03)	3.43 (0.10)	3.51 (0.11)	3.01 (0.08)*** ††	3.76 (0.25)
Microalbuminuria (mg/L)	3.46 (1.06)	7.67 (1.51)**	5.20 (1.45)	3.53 (1.39) †	7.95 (1.40)***
Five-year risk score of CVD	3.0% (0.10)	5.2% (0.24)***	5.6% (0.35)***	4.3% (0.57)*	6.2% (0.63)***

*0.01<p<0.05, **0.001<p<0.01, ***p<0.001 compared to European; †0.01<p<0.05, ††0.001<p<0.01,
†††p<0.001 compared to Samoan, = geometric mean
Sundborn et al. 2008a