Statistics New Zealand
TATAURANGA AOTEAROA

New Zealand period life tables: Methodology for 2012–14

New Zealand Government

Citation
Statistics New Zealand (2015). *New Zealand period life tables: Methodology for 2012–14.* Available from www.stats.govt.nz.

ISBN 978-0-478-42993-0 (online)

Published in July 2015 by
Statistics New Zealand
Tatauranga Aotearoa
Wellington, New Zealand

Contact
Statistics New Zealand Information Centre: info@stats.govt.nz
Phone toll-free 0508 525 525
Phone international +64 4 931 4600
www.stats.govt.nz

Contents

List of tables and figures

List of tables

List of figures

1 Purpose and summary

Purpose

New Zealand period life tables: Methodology for 2012–14 describes the methods we used to construct the New Zealand Period Life Tables for 2012–14, published on 8 May 2015, and the Subnational Period Life Tables: 2012–14, published on 28 July 2015. Mortality rates were also estimated differently from previous releases of Statistics NZ life tables with the use of more statistical modelling.

About period life tables

A period life table is a way of describing mortality conditions over a short period, generally 1–5 years. It consists of a collection of functions depicting the hypothetical deaths and life spans that a population would experience if it was exposed forever to the mortality rates for the period. For instance, a life table shows the proportion of people who would survive to age 60, and the average life expectancy at birth. The key input for a period life table is a set of estimated mortality rates for each age group and sex (Preston, Heuveline & Guillot, 2001).

Period life tables are prepared for the total New Zealand population but also for a number of smaller populations, such as the Māori ethnic group and areas within the country. Calculating life tables for small populations is challenging, because the number of observations within each age-sex cell can be very low. The statistical models that were used for the life table calculations were designed specifically for this sort of 'sparse' data.

Summary of methods

To estimate the mortality rates for the 2012–14 life tables, we used a type of statistical model called a hierarchical Bayes model. This type of model copes with sparse data by sharing information across estimates: for instance, estimates for 20-year-old females are informed by results for 20-year-old males, as well as for 19-year-old and 21-year-old females. This removes the need for analysts to intervene to manually smooth or adjust implausible values. It means that the process of estimating the mortality rates is more efficient and less dependent on the prior beliefs of the analyst. Statistical models such as a Bayes model also yield explicit measures of uncertainty, which can be important when estimates are based on very small numbers of observations.

2 Input data

Estimates of mortality rates require data on numbers of deaths and on population at risk, disaggregated by age and sex, and, depending on the life table, on ethnicity and geographic area. Although we are primarily interested in the period 2012–14, we also used data from 2005–07 to give more precise estimates.

Deaths

Data on deaths came from death registrations. The data were organised by time of registration rather than by time of occurrence. This avoided the need to wait for late registrations. In practice, for periods of a year or more, the difference is generally small since the lags are typically short.

The deaths data are highly accurate. The main possible cause for concern is the accuracy of date of birth information at the highest ages. To minimise the possibility for error, we aggregated all deaths above age 100 years at the national level and above age 90 at the subnational level. We intend to investigate age misreporting at the highest ages further.

Individual death records have address information coded to the standard meshblock code, and deaths data may then be aggregated for regional councils or territorial authorities. A small number of records have address information that is not identifiable. Values for these records were imputed based on the assumption that people who have missing information have the same geographical distribution as people of the same age and sex who have complete information.

Slightly less than 0.5 percent of records in 2012–14 had missing data for ethnicity. The missing ethnicities were also imputed. In the case of ethnicity, however, we used a process called 'multiple imputation'. See Multiple imputation for ethnicity non-response in deaths data for this description.

Population-at-risk

The life table calculations use the estimated resident population at the mid-point of the period, for example at 30 June 2013 for the 2012–14 period. The population estimates for the period mid-point represent the most recent update of the estimated resident population following the most recent census. For information about the most recent update to the base estimated resident population at 30 June see Estimated resident population 2013: Data sources and methods on the Statistics NZ website.

See Population of oldest age groups for the special methods we used to calculate the population at risk for the oldest age groups.

Ethnicity

The ethnicity concept used in life tables is the ethnic group or groups that people identify with or feel they belong to. Ethnicity is self-perceived and people can belong to more than one ethnic group. For example, people can identify with Māori ethnicity even though they may not be descended from Māori. Conversely, people may choose to not identify with Māori ethnicity even though they are descended from Māori. Ethnicity is also not the same as birthplace.

In death registrations, ethnicity is identified by the person completing the registration form and this is most likely to be the funeral director (on the advice of a family member). In the population-at-risk data, ethnic group estimates are based on individuals' responses at the most recent census. Life tables for the ethnic groups are derived from total responses to the ethnic group as recorded in deaths data and as estimated by the base population respectively.

3 Period life table publications

Table 1 provides an overview of the life tables produced for the 2012–14 period using the methodology outlined in this report. Differences in combined demographic variables of age, ethnicity, and geography are a characteristic of the available period life tables.

Table 1

Period life table specification	
Period life table	Demographic specification
Complete period life tables	Age: 0, 1, 2, …, 100+ ; Sex: Male, female, total Ethnicity: Māori, Pacific, Asian, European or Other, non-Māori
Abridged life tables	Age: 0, 1-4, 5-9, …, 100+ Sex: Male, female, total
Subnational life tables	Age intervals: 0, 1-4, 5-9, …, 90+ Sex: Male, female, total Geography: 16 regional councils, 66 territorial authorities and 21 local boards for Auckland, NZ Deprivation Area (deciles) Ethnicity: Māori and non-Māori for the 16 regional councils

4 The life table functions

We compiled the period life tables for 2012–14 using standard life table functions (Preston et al, 2001). See table 2 for an overview of how the functions were derived and their notation.

Table 2

Life table functions and notation

Notation	**Description**	**Calculation**
x	Exact age in years at start of interval	
n	Length of age interval in years. In complete period life tables *n* = 1, and in abridged life tables *n* = 1, 4 or 5.	
${}_na^x$	Average person-years lived in the interval by those dying in the interval	${}_na^x = n / 2$ for x = 1, 2, 3, … years ${}_1a^0 = 0.12$ for x = 0 years
l^x	Number left alive at exact age *x*	l^0 = 100,000 (the radix) $l^{x+n} = l^x * {}_np^x$, x = 1, 2, …
${}_nL^x$	Person-years lived between ages *x* and *x+n*	$(l^x - {}_nd^x) + {}_na^x$
${}_nd^x$	Number dying between ages *x* and *x+n*	$l^x - l^{x+n}$
${}_np^x$	Probability of surviving from age *x* to age *x+n*	$1 - {}_nq^x$
${}_nq^x$	Probability of dying between ages *x* and *x+n*	$n * {}_nm^x / (1 + (n - {}_na^x) * {}_nm^x)$
${}_nm^x$	Death rate (central death rate) in the cohort between ages *x* and *x+n*	${}_nd^x / {}_nL^x$
e^x	Expected number of years of life remaining at age *x*	$(\sum_{i=x}^{\infty} {}_nL^i) / l^i$

Note: Other equivalent derivations of life table functions are possible (Preston et al).

For cohorts aged one year or above, we assumed that, on average, people dying in the age interval do so half-way through (ie ${}_1a^x = 0.5$ and ${}_5a^x = 2.5$). Infant mortality observed in the youngest cohort (age 0 years) is more concentrated at the earlier stages of infancy. Based on detailed data from birth and death registrations, we calculated that babies who died in their first year of life died on average after 0.12 of a year (ie ${}_1a^0 = 0.12$).

5 Data-specific methods

This chapter describes the methods that were applied to the input data. In deaths data we used a multiple imputation process to estimate ethnicity indicators in records with missing responses. In population-at-risk data we used an extinct generation method to estimate the population size at the very high ages.

Multiple imputation for ethnicity non-response in deaths data

In the analysis file, there were four indicator variables measuring each person's ethnicity: an indicator for Māori ethnicity; for Pacific ethnicity; for Asian ethnicity; and for European or Other ethnicity. An individual can have a positive value for more than one ethnicity indicator – that is, an individual can have multiple ethnicities.

A small proportion of records had no value for any of the indicator variables because the ethnicity question on the registration form was left blank or could not be coded. We dealt with missing values on the ethnicity indicators by using a statistical technique called multiple imputation (Rubin, 2004).

The first step in this process was to construct a statistical model for each of the ethnicity indicators, given the probability that each individual will belong to the ethnicity, given the time period, and the individual's age, sex, and region of residence. The model used data from records where no ethnicity indicators were missing. The model is a binomial logistic regression, implemented with the glm ('generalised linear model') function in the statistical programming language R.

We then used coefficients from each model to calculate the probability that a person with missing data on ethnicity belongs to each ethnic group. For instance, the model for Māori ethnicity might suggest that a person with a particular combination of age, sex, and region of residence had an 80 percent chance of belonging to the Māori ethnic group, while the model for Pacific ethnicity might suggest that the person had a 50 percent chance of belonging to a Pacific ethnic group.

The predicted probabilities were used to randomly assign values to each person with missing data on ethnicity. The process of random assignment was done three times, yielding three filled in datasets. When mortality rates are estimated, the estimation process is carried out three times – once for each filled in dataset – and then the results are combined. Creating multiple datasets, and pooling results from multiple datasets, means that the additional uncertainty resulting from the missing ethnicity indicators is reflected in the final results.

Population of oldest age groups

Estimating size of the population at very high ages, such as 100 and over, is difficult, since very old people can be hard to capture in a census, and since even a small number of people overstating their age can have a big proportional effect on the apparent size of this population. We therefore estimated the oldest age groups using the 'extinct generation method' (Preston et al, 2001). This method exploits the fact that death registrations in a country such as New Zealand are usually highly accurate. Starting from the assumption that everyone from a cohort is likely to have died by a certain year and age, it is possible to count backwards and estimate the size of the cohort in previous years.

The 2006 population-at-risk has the last eight years of life adjusted to the extinct cohort count since we have eight years of deaths data available since the 30 June 2006 reference point (ie if the oldest age at death in 2005–07 for Māori males was 105 years, the population estimates for ages 98–105 years will be replaced with extinct cohort counts, and any non-zero population estimates above these ages will be zero). The 2013 population-at-risk will have the last year of life adjusted to the extinct cohort method, with the age adjusted uniquely for the population-at-risk group.

Consistency of ethnicity reporting in deaths and population-at-risk data

Reporting of ethnicity of the deceased by next-of-kin may not be consistent with the ethnicity reported in their latest census. For this reason, total responses to the ethnic group may be under- or over-reported in deaths data compared with the population-at-risk data. This is referred to as a numerator-denominator bias in the calculation of death rates.

Although the overall differences in total responses observed for each of the ethnic groups were small, we acknowledge there were differences when comparing relative responses across age groups and ethnic groups. Death rates estimated for the 2012–14 ethnic life tables have not been adjusted for a numerator-denominator bias.

We are investigating statistical methods for incorporating adjustments to death rates for ethnic groups as part of future releases of period life tables. Specifically we are looking at the consistency of ethnicity reporting in deaths and census data by comparing ethnicity responses in a unit record dataset of integrated deaths and 2013 Census records. The results will be available following the publication of ‘Integrating New Zealand Census with New Zealand birth and death registrations: Privacy impact assessment’, which will be available on Data integration projects.

6 Estimation of death rates

This chapter explains how we dealt with the challenge of random variation in disaggregated deaths data and it describes the hierarchical Bayesian model that was implemented for estimating death rates.

Dealing with random variation

The main challenge when estimating death rates is dealing with random variation in the number of deaths. In some age-sex groups, particularly when disaggregating by geographical area, or when dealing with a single ethnic group, only 0, 1, or 2 deaths occur in most years. This means that random variation leads to large fluctuations in the observed death rates. If an age-sex group has two deaths in one year, for instance, and one death in the next, then the observed death rate for that age-sex group declines by 50 percent. It is highly unlikely, however, that the true, underlying risk of dying has declined by 50 percent.

Demographers typically deal with randomness caused by small numbers by aggregating the data, or by drawing smooth lines through the observed rates using techniques such as splines or model life tables. The problem with aggregating is that much of the important fine detail is then lost. The problem with fitting smooth curves is that the assumption that the underlying rates are smooth may not always hold, such as in the late teens when mortality rates increase sharply. Traditional smoothing methods also require manual intervention, which increases the amount of work required, and reduces transparency and replicability.

Our approach was to estimate mortality rates using a hierarchical Bayesian model. This model contains a sub-model that accounts for the year-to-year or age-to-age random variation in observed counts, and a sub-model that describes the underlying risk of dying. The combined model smooths heavily in places where there are few observations, and hardly smooths at all in places where there are abundant observations. Gelman and Hill (2006) is an excellent general introduction to hierarchical Bayesian models. An application of hierarchical Bayesian models to estimating mortality rates in New Zealand is Blakely et al (2009).

We also addressed the problem of small sample sizes by bringing as much data to bear on the problem as possible. We used deaths data for the period 2005–07 in addition to that for 2012–14, even though we were primarily interested in results for 2012–14. The extra three years of data provided more stable estimates, provided we were willing to assume that the age-sex pattern of deaths remained similar over the period 2005–14. As discussed below, this assumption is borne out by the data.

The reason we pooled deaths over three-year periods was also to increase the number of observations available for estimating small domains. Using longer periods such as five-year periods would increase the data still further, but the meaning of the results starts to become unclear as the reference period lengthens. To measure exposure to the risk of dying we used three times the census-year estimated resident population at 30 June.

The model

Let d_{xst} be deaths of people aged *x* of sex *s* during period *t* (where *t* = 2005–07 and 2012–14). Similarly, let n_{xst} be the population-at-risk of dying, and let θ_{xst} be the underlying mortality rate, the quantity we wish to estimate. We assume that deaths follow a Poisson distribution,

$$d_{xst} \sim \text{Poisson}(\theta_{xst} n_{xst}) \,, \qquad (1)$$

which is a standard assumption in demographic modelling.

We in turn model log death rates using

$$\log \theta_{xst} \sim \text{N}(\beta^0 + \beta_x^{age} + \beta_s^{sex} + \beta_{xs}^{age:sex} + \beta_t^{time},\ \sigma^2) \,, \qquad (2)$$

where β^0 is an intercept, β_x^{age} is an age effect, β_s^{sex} is a sex effect, β_t^{time} is a time effect, and $\beta_{xs}^{age:sex}$ is an interaction between age and sex. The use of a model such as (2) is a standard way of dealing with small cell sizes. When the number of observations in an age-sex cell is small, the estimate draws heavily on the predictions from the model. This is a way of ‘borrowing strength’ from the dataset as a whole.

The intercept, sex effect, and time effect are all treated as ‘fixed effects’, that is,

$$\beta^0 \sim \text{Uniform}(-\infty, \infty)$$

$$\beta_s^{sex} \sim \text{Uniform}(-\infty, \infty)$$

$$\beta_t^{time} \sim \text{Uniform}(-\infty, \infty) \,.$$

The age effect is modelled using

$$\beta_x^{age} \sim \text{t}_4(\gamma_x^{age},\ \sigma_{age}^2)$$

$$\gamma_x^{age} \sim \text{N}(\gamma_{x-1}^{age} + \delta_{x-1}^{age},\ \tau_{age}^2)$$

$$\delta_x^{age} \sim \text{N}\left(\delta_{x-1}^{age},\ \omega_{age}^2\right) .$$

This is an integrated random walk with noise (Prado and West, 2010). The integrated random walk with noise was developed for time series, but works equally well for age groups, which have similar patterns of correlation between successive observations. The ‘integrated’ in the title refers to the fact that there is an upward trend. The ‘noise’ refers to the fact that the model distinguishes between one-off changes that affect a single age group and systematic changes which shift the overall level. The use of a Student t distribution for idiosyncratic errors gives the model extra robustness, which it needs for sudden changes such as the difference between mortality at age 0 and age 1.

The error terms, σ_{age}, τ_{age} and ω_{age} have independent non-informative uniform distributions, ie

$$\sigma_{age} \sim \text{Uniform}(0, \infty)$$

$$\tau_{age} \sim \text{Uniform}(0, \infty)$$

$$\omega_{age} \sim \text{Uniform}(0, \infty) \,.$$

The interaction effect of age and sex ($\beta_{xs}^{age:sex}$) is modelled as a random walk with noise,

$$\beta_{xs}^{age:sex} \sim \mathrm{t}_4\left(\gamma_{xs}^{age:sex},\, \sigma_{age:sex}^2\right)$$

$$\gamma_{xs}^{age:sex} \sim \mathrm{N}\left(\gamma_{x-1,s}^{age:sex},\, \tau_{age:sex}^2\right).$$

There is no need for an integrated model, as there is no systematic upward or downward trend in the interaction. Again, the error terms, $\sigma_{age:sex}$ and $\tau_{age:sex}$, are independently distributed as

$$\sigma_{age:sex} \sim \mathrm{Uniform}(0, \infty)$$

$$\tau_{age:sex} \sim \mathrm{Uniform}(0, \infty).$$

The model above was used to estimate results for the total population, and for Māori, non-Māori, Pacific, Asian, and European or Other ethnic groups. The model is fitted separately six times (or in fact 16 times, including multiple imputation), to generate the six sets of mortality estimates.

The basic model is extended slightly to give estimates for subnational areas, such as regional council areas and territorial authorities. We add an extra 'region' dimension, which we denote *r*, and model deaths using

$$d_{xsrt} \sim \mathrm{Poisson}(\theta_{xsrt} n_{xsrt}),$$

and death rates using

$$\log \theta_{xsrt} \sim \mathrm{N}\left(\beta^0 + \beta_x^{age} + \beta_s^{sex} + \beta_{xs}^{age:sex} + \beta_r^{reg} + \beta_t^{time},\, \sigma^2\right).$$

The region effect is modelled using

$$\beta_r^{reg} \sim \mathrm{t}_4\left(0,\, \tau_{reg}^2\right)$$

$$\tau_{reg} \sim \mathrm{Uniform}(0, \infty).$$

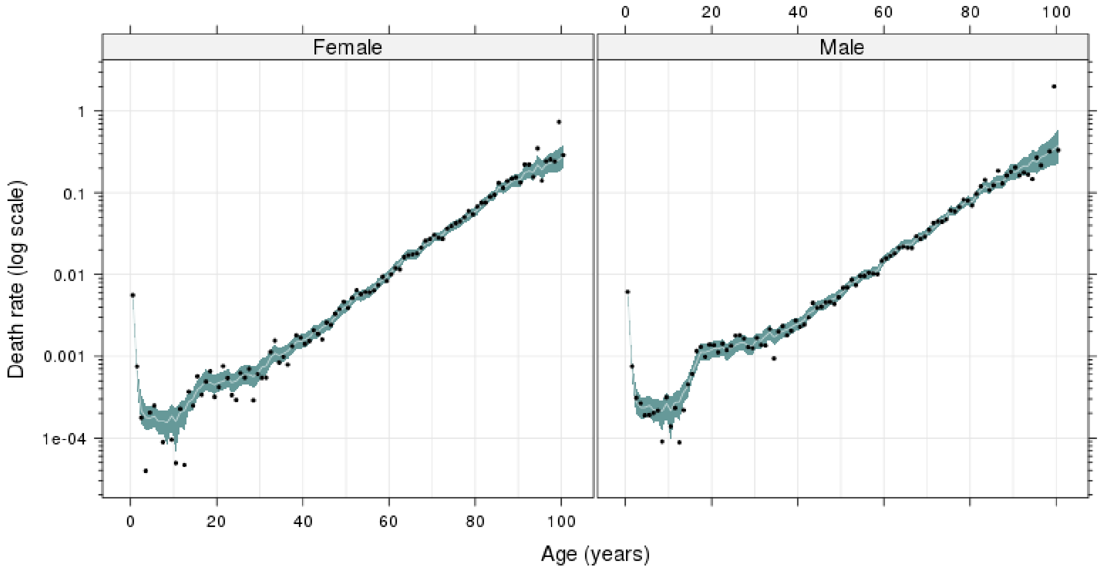
The use of a Student t distribution allows for the possibility that some regions are substantially different from the rest of the country, and avoids pulling results for unusual regions too strongly towards the national average.

The models were estimated using Markov chain Monte Carlo methods (Gelman et al, 2014). The estimation process yielded a large sample from the posterior distribution of the model. This sample can be analysed to provide a description of quantities of interest. The sample can also be used to generate distributions for derived quantities, such as the life table functions shown in chapter 4.

Figure 1 shows some typical results from the model for mortality rates. The results are for the Māori ethnic group in 2012–14. There are 101 age groups, each of a single year apart from the final 100+ age group. The green bands show 95 percent credible intervals. Under the model assumptions, there is a 95 percent chance that the true mortality rate lies within the 95 percent credible intervals. The white 'centre' line shows the median estimates. The black dots are 'direct' estimates of mortality rates, that is the actual number of deaths divided by the population at risk, with no smoothing or modelling.

The credible intervals are narrowest, indicating that uncertainty is least, around ages 60–80, where the most deaths occur. The intervals are widest around ages 2–14, where there are very few deaths. After early teens, death rates more or less follow a straight line on the log scale, aside from a period between late teens and about age 30, the 'accident hump'. Uncertainty also increases in the oldest ages, 80+, due to bigger variations in deaths counts for the lower population-at-risk.

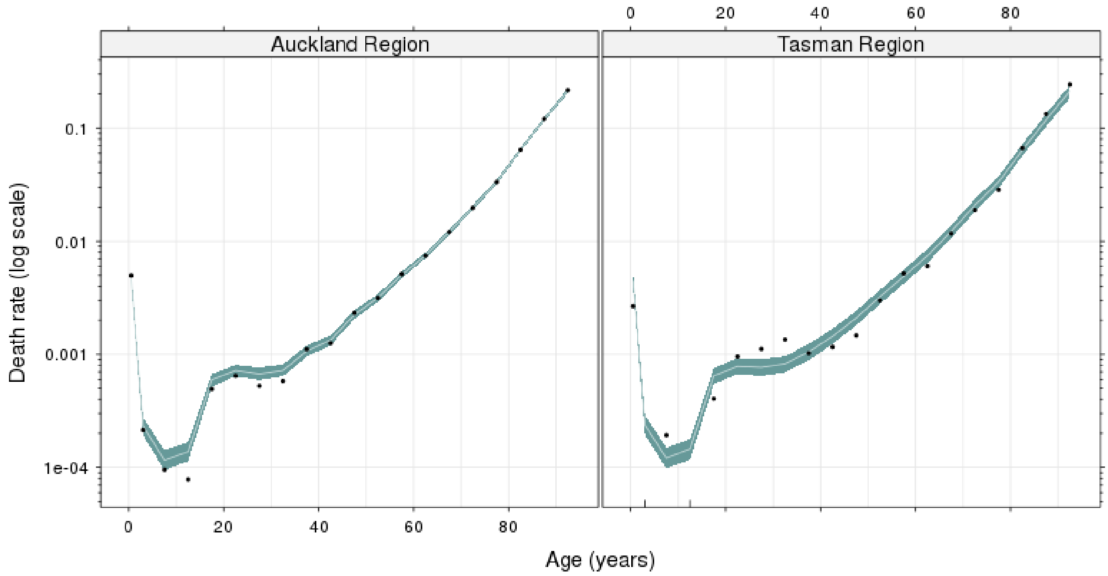
Figure 1
Death rate by age and sex
Māori ethnic group, 2012–14



Note: The green bands show 95 percent credible intervals of the modelled estimates and the white 'centre' line represents the median estimate of the death rate. The black dots are the direct estimates.
Source: Statistics New Zealand

Figure 2 shows death rates for Auckland and Tasman regions. Rates for geographic areas are estimated for age intervals (see table 1) rather than at single years. At this level of aggregation, credible intervals are narrower and the modelled results are closer to the direct estimates. The wider credible intervals for the smaller Tasman region reflects the increased random variation in death counts, including no observations for two age intervals (1–4 years and 10–14 years as marked by the opposite pointing tick marks on the axis in figure 2).

Figure 2
Death rate by age interval for males
Auckland and Tasman regions, 2012–14



Note: Age intervals for subnational areas are 0, 1-4, 5-9, …, 90+. The green bands show 95 percent credible intervals of the modelled estimates and the white 'centre' line represents the median estimate of the death rate. The black dots are the direct estimates.
Source: Statistics New Zealand

7 Comparison of 2005–07 published death rates with estimates based on the 2012–14 model

This chapter illustrates the differences between death rates using a cubic spline method with estimates using the hierarchical Bayesian model.

Methods in earlier period life tables

Life tables for the total population and ethnic groups published for 2005–07 and earlier periods were compiled from estimated death rates using a cubic spline method. Subnational life tables for 2005–07 and earlier periods were compiled from death rates estimated by the Brass Logit technique. See the Period life tables previously published in May 2009.

For the 2012–14 period estimation of death rates, data for the 2005–07 period were also input to the model (see Dealing with random variation). Subsequently, to show the validity of the Bayesian model, estimated death rates for the 2005–07 period based on the 2012–14 methods are compared with published rates for this period.

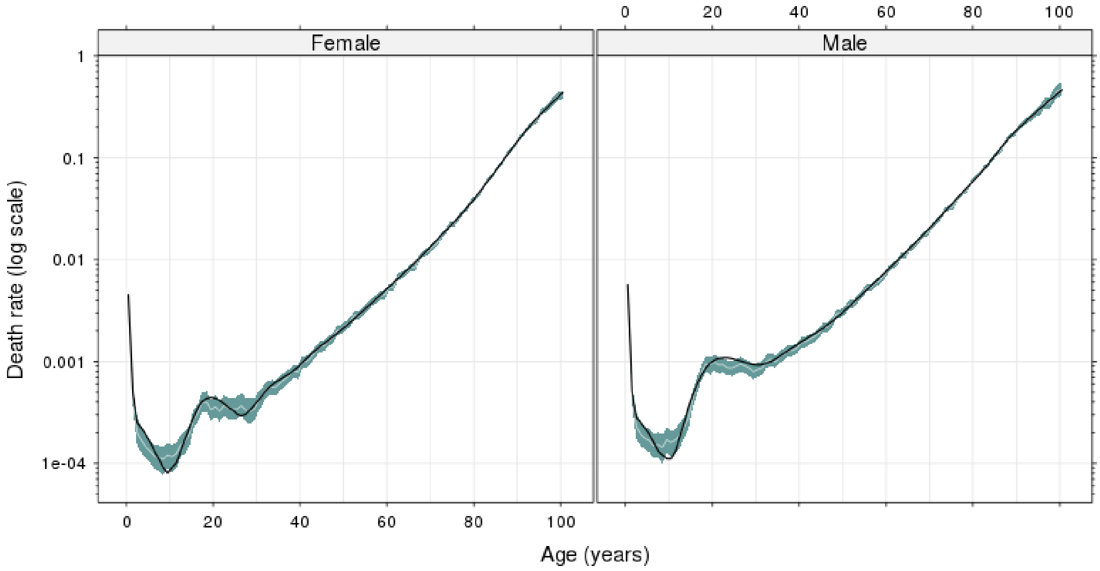
The following figures illustrate how the previously published death rates would look using the Bayesian model.

Differences in central death rates minor for total population

The overall differences in the central death rates due to the changed methods were minor for the whole population. Observed death rates for the total population are likely to have more graduated variations from one age to the next due to the larger counts of deaths at each age. In this case death rates estimated by the statistical model or a cubic spline method will be close to the direct estimate of the death rate given by the data (see figure 3).

Life expectancy at birth for females and males was 82.2 and 78.0 years respectively, as published in the 2005–07 period life tables. Equivalent figures were 82.3 and 78.3 years for the same period using the 2012–14 methodology.

Figure 3
Death rate by age and sex, total population, 2005–07
Published rates, and rates using 2012–14 methodology



Note: The black line represents the death rate using cubic spline method published for 2005–07. The green bands show 95 percent credible intervals of the modelled estimates and the white 'centre' line represents the median estimate of the death rate.
Source: Statistics New Zealand

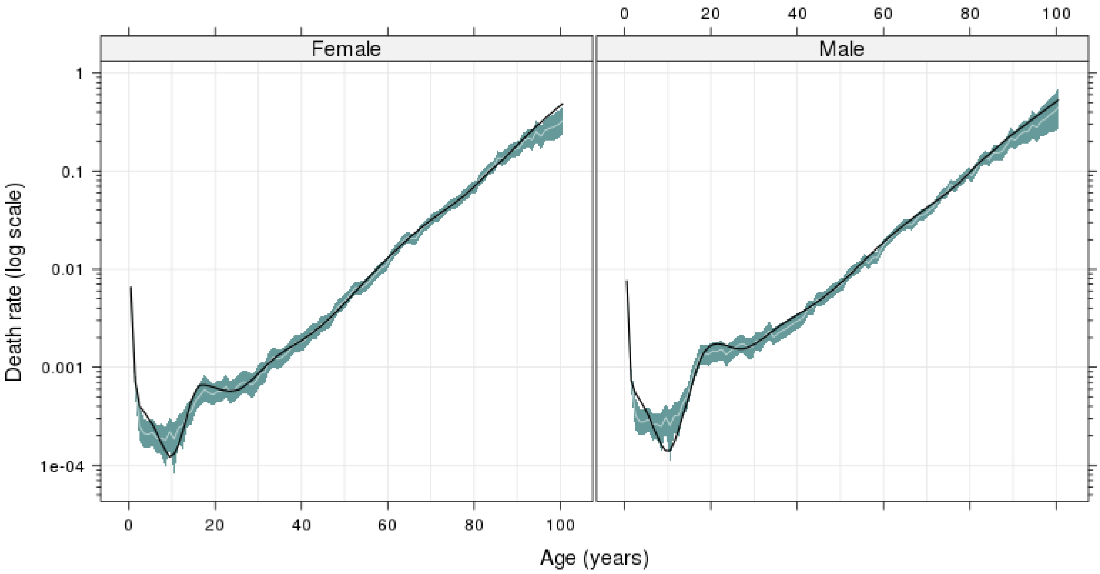
Differences in life expectancy at birth were larger for smaller populations

Differences in life expectancy at birth were larger for smaller populations, for example for Māori, due to larger variations in estimated death rates among the younger and oldest ages between the two methods (see figure 4). However, published death rates by age based on the cubic spline method were mostly within the 95 percent credible intervals provided by the 2012–14 method. The differences observed for estimated life expectancy at birth (2005–07) were larger by 0.5 years for Māori males and 0.2 years for Māori females using the statistical modelled estimates compared with the cubic spline method.

Figure 4

Death rate by age and sex, Māori population, 2005–07

Published rates, and rates using 2012–14 methodology



Note: The black line represents the death rate using cubic spline method published for 2005–07. The green bands show 95 percent credible intervals of the modelled estimates and the white ‘centre’ line represents the median estimate of the death rate.
Source: Statistics New Zealand

Published death rates for regions were within the 95 percent intervals estimated by the model

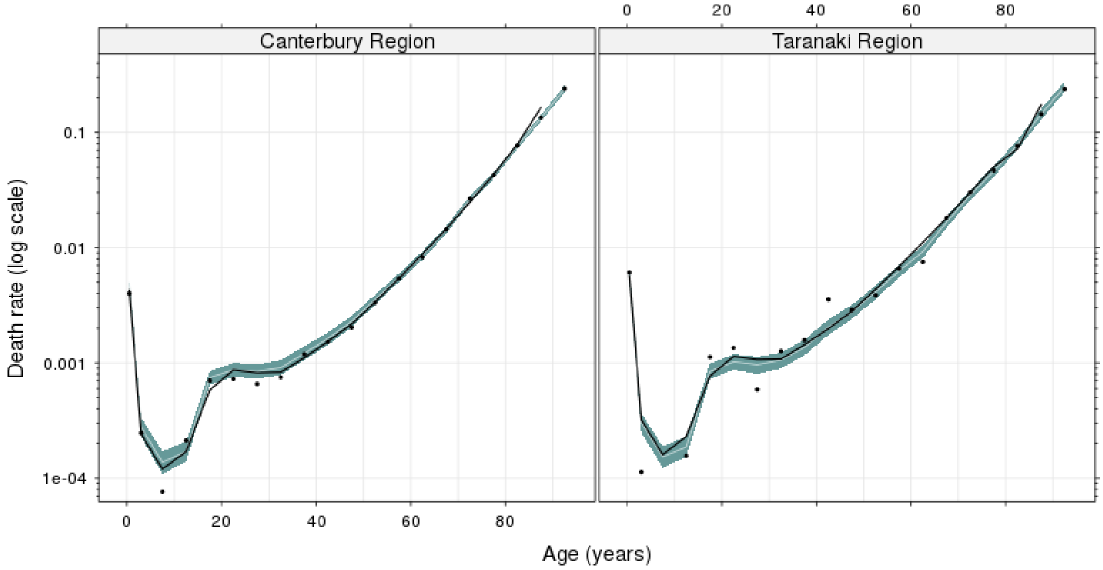
Figure 5 compares estimated death rates by age interval published for 2005–07 with the modelled estimates (2012–14 method) for two regions. Estimates for Canterbury region, with the largest population of the two, follow closely the direct estimates of the observed data, and differences between published estimates and the modelled median estimates are very small.

For the smaller Taranaki region there is increased random variation in observed death rates among youth and young adults due to low counts for these age intervals. Published death rates for this region are mostly within the bounds of the 95 percent credible intervals estimated by the model (2012–14 method). This is a general result that is applicable when comparing published death rates for regions with the new method.

Figure 5

Death rate by age interval for Canterbury and Taranaki regions, 2005–07

Published rates, and rates using 2012–14 methodology



Note: Age intervals for subnational areas are 0, 1-4, 5-9, …, 90+. The black line represents the death rate using Brass logit technique published for 2005–07. The green bands show 95 percent credible intervals of the modelled estimates and the white 'centre' line represents the median estimate of the death rate. The black dots are the direct estimates.

Source: Statistics New Zealand

Glossary

Age-specific

A measure relating to an age group.

Asian ethnicity

People who identify with an Asian ethnicity (eg Chinese, Indian, Korean) with or without other ethnicities. Because ethnicity is self-perceived, people can identify with an Asian ethnicity even though they are not descended from an Asian ancestor. Conversely, people may choose to not identify with an Asian ethnicity even though they are descended from an Asian ancestor.

Census night population count

A count of all people present in a given area on a given census night. The census night population count of New Zealand includes visitors from overseas who are counted on census night, but excludes New Zealand residents who are temporarily overseas.

For a subnational area, the count includes visitors from overseas and elsewhere in New Zealand (people who do not usually live in that area), but excludes residents of that area who are temporarily elsewhere on census night (people who usually live in that area but are absent).

Census usually resident population count

A count of all people who usually live in a given area, and are present in New Zealand, on a given census night. The census usually resident population count of New Zealand excludes visitors from overseas and excludes New Zealand residents who are temporarily overseas.

For a subnational area, the count excludes visitors from overseas and elsewhere in New Zealand (people who do not usually live in that area), but includes residents of that area who are temporarily elsewhere in New Zealand on census night (people who usually live in that area but are absent).

Cohort

A group of people sharing a common demographic experience. For example, the 1900 birth cohort refers to the people who were born in the year 1900.

Death

The permanent disappearance of all evidence of life at any time after live birth has taken place (postnatal cessation of vital functions without capability of resuscitation). This definition therefore excludes foetal deaths.

Death (mortality) rate

The number of deaths relative to the exposed-to-risk population, often expressed as a rate per 1,000 population.

Estimated de facto population

An estimate of all people present in a given area at a given date. The estimated de facto population of New Zealand includes all people present in New Zealand and counted by the census (census night population count). This estimate includes visitors from overseas who are counted on census night, but excludes New Zealand residents who are temporarily overseas.

For a subnational area, the estimate includes visitors from overseas and elsewhere in New Zealand (people who do not usually live in that area), but excludes residents of that area who are temporarily elsewhere on census night (people who usually live in that area but are absent).

The estimated de facto population at a given date after census includes births, deaths and net migration (arrivals less departures) of people during the period between census night and the given date.

De facto population estimates are no longer produced. National population estimates were produced annually (reference date at 31 December) from 1936 to 1950, and quarterly (reference dates at 31 March, 30 June, 30 September, and 31 December) from March 1951 to June 1997. Subnational population estimates were produced annually (reference date at 31 March) to 1995.

Estimated resident population

An estimate of all people who usually live in a given area at a given date. The estimated resident population of New Zealand includes all residents present in New Zealand and counted by the census (census usually resident population count), residents who are temporarily overseas (who are not included in the census), and an adjustment for residents missed or counted more than once by the census (net census undercount). Visitors from overseas are excluded.

For a subnational area, the estimate excludes visitors from overseas and elsewhere in New Zealand (people who do not usually live in that area), but includes residents of that area who are temporarily elsewhere on census night (people who usually live in that area but are absent).

The estimated resident population at a given date after census includes births, deaths and net migration (arrivals less departures) of residents during the period between census night and the given date.

National population estimates are produced quarterly (reference dates at 31 March, 30 June, 30 September, and 31 December) and subnational population estimates are produced annually (reference date at 30 June).

Ethnicity

Ethnicity is the ethnic group or groups that people identify with or feel they belong to. Ethnicity is a measure of cultural affiliation, as opposed to race, ancestry, nationality, or citizenship. Ethnicity is self-perceived and people can belong to more than one ethnic group.

An ethnic group is made up of people who have some or all of the following characteristics:

- a common proper name
- one or more elements of common culture which need not be specified, but may include religion, customs or language
- unique community of interests, feelings and actions
- a shared sense of common origins or ancestry
- a common geographic origin.

This definition is based on the work of A Smith (1986), “The Ethnic Origins of Nations”.

People can identify with an ethnicity even though they are not descended from ancestors with that ethnicity. Conversely, people may choose to not identify with an ethnicity even though they are descended from ancestors with that ethnicity. Ethnicity is not the same as birthplace.

In the Census of Population and Dwellings, ethnicity is identified by the person completing the census form. In the case of births and deaths, ethnicity is identified by the person completing the registration form. For births this is usually the parent(s), while for deaths this is most likely to be the funeral director (on the advice of a family member).

For more information about ethnicity, refer to the Review of the Measurement of Ethnicity on the Statistics NZ website which includes information about the Statistical Standard for Ethnicity 2005.

European or Other ethnicity

People who identify with a European ethnicity (eg New Zealand European, English, Dutch) or Other ethnicity (eg New Zealander), with or without other ethnicities. Because ethnicity is self-perceived, people can identify with a European ethnicity even though they are not descended from a European ancestor. Conversely, people may choose to not identify with a European ethnicity even though they are descended from a European ancestor.

Life tables have been compiled for the 'European or Other (including New Zealander)' ethnic group. This reflects that sufficient demographic data is available to enable estimation of death rates for the combined ethnic grouping, but not for the separate ethnic groups. This approach is consistent with Guidelines for Using Ethnicity Data: 2006 Census, available on the Statistics NZ website.

Exposed-to-risk population

People able to experience a particular event (eg death) often because of specific characteristics (eg age, sex, geographic location).

Foetal death

Death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of gestation, by either spontaneous abortion (miscarriage), induced abortion, or stillbirth. Death is indicated by the fact that after such separation the foetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles.

Infant death

The death of a child (who was born alive) before the age of one year.

Life expectancy

The average length of life remaining at a given age. In a period life table, it is the average length of life from a given age, assuming people experience the age-specific mortality rates of a given period from the given age onwards. In a cohort life table, it is the average length of life from a given age, of people born in a given period, based on the mortality rates actually experienced by them from that given age onwards.

Life expectancy at birth

The average length of life of a birth cohort. In a period life table, it is the average length of life of newborn babies, assuming they experience the age-specific mortality rates of a given period throughout their life. In a cohort life table, it is the average length of life of people born in a given period, based on the mortality rates actually experienced by them throughout their life.

Life table

A tabular numerical representation of mortality and survivorship of a cohort of births at each age. It comprises an array of measures, including probabilities of death, probabilities of survival, and life expectancies at various ages.

Complete life tables present life table functions for each single year of age, while abridged life tables present life table functions for age groups.

Current, period, or cross-sectional life tables are based on current mortality rates. These tables assume that as a cohort passes through life it experiences a given pattern of age-specific mortality rates, which do not change from year to year. Although it is usually based on death rates from a real population during a particular period of time, these tables are a hypothetical model of mortality as they do not describe the real mortality that characterises a cohort as it ages.

Cohort, longitudinal, or generation life tables are based on the actual mortality experience of a particular cohort (eg all people born in the year 1900). These tables require data over many years, from infancy to the oldest age lived by the cohort (ie until the death of the last survivor).

Live birth

The birth of a child who breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached. Each product of such a birth is considered liveborn. All liveborn infants should be registered and counted as such, irrespective of gestation or whether alive or dead at the time of registration. If they die at any time following birth they should also be registered and counted as deaths.

Longevity

Length of life.

Māori ethnicity

People who identify with the Māori ethnicity with or without other ethnicities. Because ethnicity is self-perceived, people can identify with Māori ethnicity even though they are not descended from a Māori ancestor. Conversely, people may choose to not identify with Māori ethnicity even though they are descended from a Māori ancestor.

Mean population

The average number of people in an area during a given period, usually a year. This measure may be estimated in terms of a simple or weighted arithmetic mean of monthly or quarterly population during the reference period. If the mean population is unavailable, the population at the midpoint of the period is generally suitable for most purposes.

Mortality

The death of individuals in a community.

Net census undercount

The difference between undercount and overcount. It is usually expressed as a percentage of what should have been the complete count rather than as a percentage of what was counted. The 1996, 2001, and 2006 post-enumeration surveys estimated net census undercount to be 1.6 ± 0.2, 2.2 ± 0.3, and 2.0 ± 0.4 percent, respectively (sample errors at the 95 percent level).

Non-Māori ethnicity

People identifying with an ethnicity excluding those who identified with Māori ethnicity.

Pacific ethnicity

People who identify with a Pacific ethnicity (eg Samoan, Tongan, Fijian) with or without other ethnicities. Because ethnicity is self-perceived, people can identify with a Pacific ethnicity even though they are not descended from a Pacific ancestor. Conversely, people may choose to not identify with a Pacific ethnicity even though they are descended from a Pacific ancestor.

Post-enumeration survey

A sample survey to check the accuracy of coverage and/or response of another census or survey. A post-enumeration survey was conducted after each of the 1996, 2001, and 2006 Censuses of Population and Dwellings to check the coverage of each census.

Radix

The original size of the birth cohort of a life table, usually set at 100,000 for convenience.

Resident population concept

A statistical basis for a population in terms of those who usually live in a given area at a given time. The census usually resident population count is a census measure of the resident population concept, and the estimated resident population is a demographic measure of the resident population concept. In terms of vital statistics, the resident population concept refers to events that relate to residents of New Zealand only.

Resident temporarily overseas

A person who usually lives in New Zealand but who is overseas for a period of less than 12 months. In census statistics, a resident temporarily overseas is a person who is identified on the census dwelling form as usually living in that dwelling but who is overseas for a period of less than 12 months. In international travel and migration statistics, a resident temporarily overseas is someone who is mainly living in New Zealand for 12 months or more, who is overseas for a period of less than 12 months.

Standardised death rates

The overall death rate that would have prevailed in a standard population if it had experienced the age-specific (usually age-and-sex-specific) death rates of the population or area being studied.

Subnational

Geographical units of a country (eg area units, territorial authority areas, regional council areas, urban areas) whose boundaries are defined for administrative, legal, or statistical purposes.

References

Ajwani, S, Blakely, T, Robson, B, Tobias, M, & Bonne, M (2003). *Decades of disparity: Ethnic mortality trends in New Zealand 1980–1999*. Wellington: Ministry of Health and University of Otago. Available from www.health.govt.nz.

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

Gelman, A, & Hill, J (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A, Carlin, JB, Stern, HS, Dunson, DB, Vehtari, A, & Rubin, DB (2014). *Bayesian Data Analysis*. CRC Press.

Prado, R, & West, M (2010). *Time Series: Modeling, Computation, and Inference.* USA: Chapman & Hall/CRC Texts in Statistical Science.

Preston, SH, Heuveline, P, & Guillot, M (2001). *Demography: Measuring and modeling population processes*. Oxford, United Kingdom: Blackwell Publishers Ltd.

Rubin, DB (2004). Multiple imputation for nonresponse in surveys. John Wiley & Sons.

Statistics New Zealand (2015). Data integration projects. Available from www.stats.govt.nz.

Statistics New Zealand (2014). Estimated resident population 2013: Data sources and methods. Available from www.stats.govt.nz.

Statistics New Zealand (2015). Period life tables. Available from www.stats.govt.nz.

Venables, WN, & Ripley, BD (2002). *Modern Applied Statistics with S* (4th ed). New York: Springer.

Appendix: Historic overview of NZ period life tables 1950–52 to 2012–14

Summary of ethnic concepts, deaths measures, population measures, life table changes and revisions

2012–14

Ethnic concept

- Māori life tables based on deaths of people with Māori ethnicity and the population of Māori ethnicity.
- Non-Māori life tables based on deaths and population not included in Māori life tables.
- Pacific life tables based on deaths of people with Pacific ethnic group and the population of Pacific ethnic group.
- All ethnic life tables make allowance for deaths and population with no ethnic response.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand, by geographic area.

Population measure

- Estimated resident population at 30 June 2013 (based on census usually resident population count at 5 March 2013), by geographic area.

Life table changes and revisions

- National Pacific life tables included in the release, provided back to 2005–07.
- National European or other, and Asian life tables made available back to 2005–07.
- Subnational life tables extended to include all territorial authorities, available back until 1995–2007.
- Release included Māori and non-Māori life tables by regional council and life tables by deprivation index back until 2005–07.
- Bayesian statistical modelling is introduced.
- This did not create a discontinuity in national life tables.
- It created some differences in national ethnic life table results, but these weren't significant enough to revise previous periods.
- It created a discontinuity in subnational time series. Series was revised back to 1995–97.

2005–07

Ethnic concept

- Māori life tables based on deaths of people with Māori ethnicity and the population of Māori ethnicity.
- Non-Māori life tables based on deaths and population not included in Māori life tables.

- Both Māori and non-Māori life tables make allowance for deaths and population with no ethnic response.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand.

Population measure

- Estimated resident population at 30 June 2006 (based on census usually resident population count at 7 March 2006).

Life table changes and revisions

- Cubic spline method used for national and ethnic life tables in 2005–07 and prior.
- Brass-logit method used for subnational life tables in 2005–07 and earlier periods.

2000–02

Ethnic concept

- Māori life tables based on deaths of people with Māori ethnicity and the population of Māori ethnicity.
- Non-Māori life tables based on deaths and population not included in Māori life tables.
- Both Māori and non-Māori life tables make allowance for deaths and population with no ethnic response.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand.

Population measure

- Estimated resident population at 30 June 2001 (based on census usually resident population count at 6 March 2001).

1995–97 revised

Ethnic concept

- Māori life tables based on deaths of people with Māori ethnicity and the population of Māori ethnicity.
- Non-Māori life tables based on deaths and population not included in Māori life tables.
- Both Māori and non-Māori life tables make allowance for deaths and population with no ethnic response.
- Adjustment made to Māori/non-Māori deaths, for undercount of Māori deaths using adjustment ratios presented in Ajwani (2003).

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand.

Population measure

- Estimated resident population at 30 June 1996 revised (based on census usually resident population count at 5 March 1996).

1995–97

Ethnic concept

- Māori life tables based on deaths of people with Māori ethnicity and the population of Māori ethnicity.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no ethnic response.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand.

Population measure

- Estimated resident population at 30 June 1996 (based on census usually resident population count at 5 March 1996).

1990–92

Ethnic concept

- Māori life tables based on deaths of people with any degree of Māori blood and the population of Māori ethnicity.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths with no indication of degree of Māori blood, and population with no ethnic response.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Census usually resident population count at 5 March 1991.

1985–87

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood' and the population of 'single ethnic response' Māori.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths with no indication of degree of Māori blood and including population with no ethnic response.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1985–87 (based on census night population count at 4 March 1986).

1980–82 revised

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood' and the population of 'half or more Māori blood'.

- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.
- No adjustment was made to Māori/non-Māori deaths for undercount of Māori deaths.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1980–82 (based on census night population count at 24 March 1981).

1980–82

Ethnic concept

- Māori life tables based on deaths of people with any degree of Māori blood and the population of 'half or more Māori blood'.
- Deaths with any degree of Māori blood used as proxy adjustment for undercount of Māori 'half or more' deaths.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1980–82 (based on census night population count at 24 March 1981).

1975–77

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood' and the population of 'half or more Māori blood'.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1975–77 (based on census night population count at 23 March 1976).

1970–72

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood' and the population of 'half or more Māori blood'.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1970–72 (based on census night population count at 23 March 1971).

1965–67

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood' and the population of 'half or more Māori blood'.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1965–67 (based on census night population count at 22 March 1966).

1960–62

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood provided that the remaining blood is European or Polynesian', and the population of 'half or more Māori blood provided that the remaining blood is European or Polynesian'.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population mean year ended 31 December 1960–62 (based on census night population count at 18 April 1961).

1955–57

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood provided that the remaining blood is European or Polynesian', and the population of 'half or more Māori blood provided that the remaining blood is European or Polynesian'.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population at 30 June 1956 (based on census night population count at 17 April 1956).

1950–52

Ethnic concept

- Māori life tables based on deaths of people with 'half or more Māori blood provided that the remaining blood is European or Polynesian', and the population of 'half or more Māori blood provided that the remaining blood is European or Polynesian'.
- Non-Māori life tables based on deaths and population not included in Māori life tables, including deaths and population with no indication of degree of Māori blood.

Deaths measure

- Deaths registered in New Zealand of people resident in New Zealand and people visiting from overseas.

Population measure

- Estimated de facto population at 30 June 1951 (based on census night population count at 17 April 1951).