Methodology Report 2013/14

New Zealand Health Survey

New Zealand Health Survey

Authors

This report was written by Deepa Weerasekera, Robert Templeton, Marie Ditchburn, Steven Johnston (Health and Disability Intelligence Group, Ministry of Health), Robert Clark (National Institute for Applied Statistics Research Australia, University of Wollongong, Australia) and Barry Gribben, Carol Boustead and Neil Tee (CBG Health Research Ltd).

Acknowledgements

Thank you to the many thousands of New Zealanders who gave their time to participate in the New Zealand Health Survey, and to the interviewers who worked so diligently to collect the data.

Please refer to the Ministry of Health's publication *Annual Update of Key Results 2013/14: New Zealand Health Survey* for further acknowledgements (Ministry of Health 2014a).

Contents

List of tables

List of figures

Section 1: Introduction

Overview

The New Zealand Health Survey (NZHS) is an important data collection tool, which is used to monitor population health and provide supporting evidence for health policy and strategy development. The Health and Disability Intelligence Group within the Ministry of Health's Policy business unit is responsible for designing, analysing and reporting on the NZHS. The NZHS field work is contracted out to a specialist survey provider, CBG Health Research Ltd.

The NZHS collects information that cannot be obtained more effectively or efficiently through other means, such as by analyses of hospital administrative records, disease registries or epidemiological research. For most topics in the NZHS the survey is the best source of information at a population level.

Previous New Zealand Health Surveys were conducted in 1992/93, 1996/97, 2002/03 and 2006/07. In addition, separate stand-alone surveys on specific subjects have been conducted once every three or four years as part of the wider health survey programme. These surveys covered adult and child nutrition; tobacco, alcohol and drug use; mental health; and oral health. From July 2011 all the above surveys have been integrated into the single NZHS, which is now in continuous operation.

From 2013 onwards a number of key outputs from the NZHS became Tier 1 statistics (a portfolio of the most important official statistics, essential to understanding how well New Zealand is performing). This year the four Tier 1 statistics from the NZHS are current smoking, obesity, self-rated health, and mental health status.

This NZHS methodology report outlines the procedures and protocols followed to ensure the NZHS produces the high-quality and robust data expected of official statistics (Statistics New Zealand 2007). The information specific to the current year (2013/14) of the continuous NZHS (data-collected between July 2013 and June 2014) is included in Section 8 of this report. The corresponding information for 2011/12 and 2012/13 of the NZHS can be found in the previous reports, *New Zealand Health Survey Methodology Report 2012* (Ministry of Health 2012) and *New Zealand Health Survey Methodology Report 2012/13* (Ministry of Health 2013).

Background

The NZHS forms part of the Programme of Official Social Statistics. Statistics New Zealand established this programme to develop and coordinate official social statistics across the government. As a signatory of the Protocols of Official Statistics (Statistics New Zealand 1998), the Ministry of Health employs best-practice survey techniques to produce high-quality information through the NZHS. It uses standard frameworks and classifications, with validated questions where possible, so that NZHS data can be integrated with data from other sources.

The goal of the continuous NZHS is to support the formulation and evaluation of policy by providing timely, reliable and relevant health information. This information covers population health, health risk and protective factors, and health service utilisation.

To achieve this goal a number of specific objectives have been identified. The *Content Guide 2013/14: New Zealand Health Survey* contains further information on these objectives (Ministry of Health 2014b).

The NZHS has been carefully designed to minimise the impact on respondents. Features aimed at achieving this include:

- selecting only one eligible adult and one eligible child per dwelling
- using well-tested and well-proven questionnaires
- using professional, trained interviewers to conduct the interviews
- making an appointment to conduct each interview at a time that suits the respondent and their family
- having the option of using a proxy respondent where participants living in private dwellings have severe ill health or cognitive disability.

The Multi-region Ethics Committee provided approval of the NZHS 2013/14 (Multi-region Ethics Committee Reference: MEC/10/10/103).

Section 2: The New Zealand Health Survey questionnaire

The NZHS comprises a set of core questions combined with a flexible programme of rotating topic modules. The questionnaire is administered (face to face and computer assisted) to adults aged 15 years and older, as well as to children aged 0 to 14 years, generally through their primary caregiver, who acts as a proxy respondent.

Over previous years the content of health surveys has remained similar so that data can be compared over time. The current NZHS maintains continuity with the previous surveys by including a set of core questions in both the adult and child questionnaires. The module topics change every 12 months.

Cognitive testing is undertaken to ensure the questions are understood as intended and response options are appropriate.

For more detail on the rationale of topic inclusion, cognitive testing and the content of the questionnaires, see the *Content Guide 2013/14: New Zealand Health Survey* (Ministry of Health 2014b).

Core component

The core questions for both adults and children are largely drawn from the main topic areas included in the 2006/07 NZHS and 2011/12 NZHS. Topics include long-term conditions, health service utilisation and patient experience, health risk and protective factors, health status, and sociodemographics. Table 1 summarises the topics included in the core component of the NZHS.

Table 1: Core content of the NZHS

Domain	Topics
Children	
Long-term conditions	Asthma, eczema, diabetes, rheumatic heart disease, mental health conditions
Health status and development	General health
Health behaviours	Breastfeeding, nutrition, physical activity, family cohesion
Health service utilisation and patient experience	Primary health care provider use, general practitioners, nurses, medical specialists, oral health care professionals, other health care professionals, hospital use, prescriptions
Sociodemographics	Child: sex, age, ethnicity, language, country of birth
	Primary caregiver/proxy respondent: relationship to child, age, education, income and income sources, employment status, and household characteristics
Anthropometry	Height, weight and waist circumference measurements

Domain	Topics
Adults	
Long-term conditions (self-reported)	Heart disease, stroke, diabetes, asthma, arthritis, mental health conditions, chronic pain, high blood pressure, high blood cholesterol
Health status	General health (physical and mental health), psychological distress
Health behaviours	Physical activity, tobacco smoking, vegetable and fruit intake, alcohol use and hazardous drinking
Health service utilisation and patient experience	Primary health care provider use, general practitioners, nurses, medical specialists, oral health care professionals, other health care professionals, hospital use, prescriptions
Sociodemographics	Sex, age, ethnicity, language, country of birth, education, income and income sources, employment status, medical insurance, household characteristics
Anthropometry	Height, weight and waist circumference measurements; blood pressure

Module component

All the module topics for the continuous NZHS until 2015/16 are summarised in Table 2.

Table 2: NZHS module topics, 2011/12 to 2015/16

Year of NZHS	Adult module topic(s)	Child module topic(s)
2011/12	Health service utilisation Patient experience Problem gambling Discrimination	Health service utilisation Patient experience
2012/13	Alcohol use Tobacco use Drug use	Child development Food security Exposure to second-hand smoke
2013/14	Long-term conditions Health status Disability status Living standards Housing quality	Long-term conditions Health status Disability status Living standards Housing quality
2014/15	Sexual and reproductive health	Child development
2015/16	Tobacco use	Child development

Section 3: Survey population and sample design

This section describes the target population, the survey population and the sample design for the NZHS.

Target and survey population

The *target population* is the population the survey aims to represent. The *survey population* is the population that was covered in the survey.

Target population

The target population for the NZHS is the New Zealand usually resident population of all ages (including those living in non-private accommodation). It includes those living in aged-care facilities and those temporarily living away from the household in student accommodation.

The target population is approximately 3.6 million adults (aged 15 years and over) and 0.9 million children (aged from birth to 14 years), according to the Statistics New Zealand Census population figures for 2013.

Previously (2006/07 and before) the NZHS included only people living in private accommodation. The target population for the current NZHS includes people living in non-private accommodation to improve coverage of older people in an ageing population.

Survey population

Approximately 98 percent of the New Zealand resident population of all ages are eligible to participate in the NZHS. For practical reasons a small number of households in the defined target population are excluded from the survey population. Exclusions from the survey population are:

- specific types of non-private dwellings (prisons, hospitals, hospices, dementia care units, and hospital-level care in aged-care institutions)
- households in remote areas, including areas (meshblocks) with fewer than nine occupied dwellings and those located off the main islands of New Zealand.

Sample design

The sample design for the NZHS has been developed by the National Institute for Applied Statistics Research Australia (NIASRA), University of Wollongong, Australia. For more details on how the sample size was determined and the sample design for the first three years of the survey, see *The New Zealand Health Survey: Sample design, years 1–3 (2011–2013)* (Ministry of Health 2011). For an overview of the sample design, see Clark et al 2013, available at www.pophealthmetrics.com/content/11/1/25.

Sample selection

The NZHS has a multi-stage, stratified, probability-proportional-to-size (PPS) sampling design. The survey is designed to yield an annual sample size of approximately 13,000 adults and 4500 children.

A dual frame approach has been used, whereby participants are selected from an area-based sample and a list-based electoral roll sample. The aim of this approach is to increase the sample sizes for Māori, Pacific and Asian ethnic groups.

Area-based sample

Meshblocks are the primary sampling units for the area-based sample. The geography and Census data for these meshblocks are readily available and have been used in previous New Zealand Health Surveys. The area-based sample is targeted at the ethnic groups of interest by assigning higher probabilities of selection to areas (meshblocks) in which these groups are more concentrated.

Meshblocks vary considerably in size and are therefore selected by PPS design. Through the PPS approach, larger meshblocks have a higher chance of being selected for the sample. This approach is then modified to give higher probabilities for households in areas where Māori, Pacific or Asian people are more prevalent.

A three-stage selection process is used to achieve the area-based sample. First, a sample of area meshblocks is selected within each district health board (DHB) area. Each meshblock is assigned a quarter (of the year) in which it will be surveyed.

Second, a list of households is compiled for each selected meshblock. A systematic sample of approximately 20 households is selected from this list by choosing a random start and selecting every kth household. The skip k is calculated by dividing the frame occupied dwellings count by 20.

Finally, one adult (aged 15 years or over) and one child (aged from birth to 14 years, if any in the household) are selected at random from each selected household.

Electoral roll sample

The electoral roll is another sample frame used to increase the sample size of the Māori ethnic group. The electoral roll is used to select a sample of addresses where a person has self-identified as having Māori ancestry. A copy of the electoral roll is obtained quarterly for this purpose.

Stratified three-stage sampling is used to select the sample from the electoral roll. The first stage involves selecting a sample of meshblocks within each stratum (DHB), with probability proportional to the number of addresses on the electoral roll in the meshblock. The second stage involves selecting a random sample of 10 addresses (from the list of households where any person has self-identified as having Māori ancestry) from each selected meshblock (or all addresses, if fewer than 10). The sample of meshblocks is selected so that it does not overlap with the sample from the area-based sample. Finally, one adult (aged 15 years or over) and one child (aged from birth to 14 years, if any in the household) are selected at random from each selected address.

The electoral roll is used in order to increase the recruitment rate of Māori into the sample. However, the process of contacting households and selecting an adult and child is exactly the same as for the area-based sample. In particular, the adult and child (if any in the household) randomly selected into the survey can be Māori or non-Māori. This approach ensures that probabilities of selection can be correctly calculated for all respondents.

Section 4: Data collection

Data for the NZHS are collected by CBG Health Research Ltd (CBG). The CBG interview team consists of approximately 35 professional social research interviewers. Interviews are conducted in participants' homes, with the interviewer entering responses directly into a laptop computer using the Survey System computer-assisted personal interview (CAPI) software. Show cards with predetermined response categories are used to assist respondents, where appropriate.

Dress rehearsal

A dress rehearsal was carried out in 20 meshblocks in Northland, Auckland and Waikato prior to the data collection for the 2011/12 NZHS. Its purpose was to test the sample design and to refine the instruments, operations and processes. No substantive changes were made to the sample design or instruments following the dress rehearsal.

Enumeration

CBG identifies households from meshblocks selected for the survey using the NZ Post address database, which is obtained quarterly. Each area meshblock visited by an interviewer is re-enumerated in order to record new dwellings built and those removed since the last Census enumeration and release of the NZ Post address list. The details of new dwellings are entered into CBG's Sample Manager software while the interviewer is in the field, allowing these households to be included in the random selection process of the meshblock.

Invitation to participate

The NZHS is voluntary, relying on the good will of participants, and consent is obtained without coercion or inducement. CBG uses the NZ Post address database to post each household an invitation letter from the Ministry of Health, along with an information pamphlet about the NZHS. Interviewers take copies of the information pamphlet in 11 languages when they subsequently visit households to seek people's agreement to participate in the survey.

One adult and one child (if any in the household) are randomly selected to take part in the survey using CBG's Sample Manager software. Participants are asked to sign an electronic consent form and are given a copy of the consent form to keep. The consent form includes a request for an interpreter, if required (in any of a range of different languages), and attempts are made to match respondents and interviewers by ethnicity and sex when requested.

Child interviews are conducted with a guardian/primary caregiver of the child; that is, a person who has day-to-day responsibility for the care of the child.

All participants in the NZHS are given a thank you card and a small token of appreciation, such as a pen or fridge magnet, at the conclusion of the interview. The card contains a list of health and community organisations with freephone numbers that participants can use if they would like to discuss any issues raised by their participation in the NZHS, or if they need advice on a health issue.

Call pattern

Up to 10 calls to each sampled dwelling are made at different times of the day and on different days of the week before accepting that a dwelling is a non-contact. Calls are recorded as unique events only if they are made at least two hours apart.

The number of calls made by an interviewer is spaced over two to three months. Up to six calls are made by the interviewer during the month in which the meshblock is first being surveyed. If contact with the household is not established during this first month, there is a pause for three to four weeks before attempting two more calls. Finally, there is a pause for a further three to four weeks before attempting the final two calls. This procedure helps to contact not only people who are temporarily away, but also those who are busy with work, family or socially when their dwelling is first approached.

Auditing of interviewers

CBG conducts audit calls with around 10 percent of all participants and at least one household per meshblock. Participants are also left with feedback postcards, which they can use to send feedback directly to CBG, anonymously if they choose.

Interviewer training

Interviewers take part in ongoing training courses run by CBG on how to conduct interviews.

Objective measurements

All participants aged two years and over have their height and weight measurements taken by the interviewer at the end of the survey. Those aged five years and over also have their waist circumference measured. Adults aged 15 years and over are also asked to provide a blood pressure reading. Pregnant women are excluded from having any measurements taken.

In the 2012/13 NZHS, laser height measurement was introduced. The measuring device consists of a professional laser meter (Precaster CA770) mounted to a rigid headboard, which is held against the corner of a wall or door by the interviewer. The headboard is lowered until it reaches the participant's head, whereupon the laser is activated to take a measurement. The laser design was trialled and refined in early 2012 prior to its survey-wide implementation in July 2012. The laser meter replaced traditional stadiometers, which were used in the 2011/12 NZHS.

Weight is measured with electronic weighing scales (Tanita HD-351). Participants are required to empty their pockets as well as remove their shoes and any bulky clothing that could produce an inaccurate reading.

Waist circumference is measured using an anthropometric measuring tape (Lufkin W606PM). The measurement is taken over one layer of clothing at the midpoint between the lowest palpable rib and the top of the hip bone.

Height, weight and waist circumference measurements are taken a minimum of twice each. If there is more than a 1 percent variation between the first and second measurement, then a third measurement is taken for accuracy. To align with international standards, the final height, weight and waist measurements were calculated for each respondent by taking the mean of the closest two measurements.

Blood pressure measurement was introduced to the NZHS in July 2012 and is obtained using a portable electronic sphygmomanometer (Omron HEM 907). A fabric cuff is wrapped around the participant's upper left arm, just above the elbow. Within the cuff is a plastic bladder connected by a tube to the main device. As the bladder inflates, the device is able to detect the blood pressure of the participant. The device is programmed to take three readings, with a one-minute pause between each.

Participants are left with a measurement card, which details the readings taken on the day of the survey. The card also provides information on interpreting the blood pressure measurement and where the participant can go for further information or advice.

Section 5: Response and coverage rates

The response rate is a measure of how many people who were selected to take part in the survey actually participated. A high response rate means that the survey results are more representative of the New Zealand population.

> In 2013/14, the final weighted response rates were 80 percent for adults and 85 percent for children.
>
> For more details on the response rate for 2013/14, see Section 8.

The response rate is an important measure of the quality of a survey. Methods used to maximise response rates are to:

- give interviewers initial and ongoing training and development
- support and assess interviewers in the field
- use well-designed call pattern processes, allowing for up to 10 calls to potential participants at differing times of the week and day
- revisit 'closed' meshblocks during a mop-up phase (ie, when visiting households where no contact has been established or the selected respondent was unable to take part at that time but did not refuse to participate).

Calculation of response rate

The weighted response rate reflects the probability of the household being selected into the sample. It describes the success of the survey in terms of achieving cooperation from the population being measured.

There are four components to the weighted adult response rate calculation:

1. ineligibles (eg, vacant sections, vacant dwellings and non-residential dwellings)
2. eligible responding (interview conducted, respondent confirmed to be eligible for the survey)
3. eligible non-responding (interview not conducted, but enough information collected to indicate that the household did contain an eligible adult; almost all refusals were in this category)
4. unknown eligibility (eg, non-contacts and refusals who provide insufficient information to determine eligibility).

The response rate is calculated as follows:

$$\text{Response rate} = \frac{\text{number of eligible responding}}{\left[\begin{array}{c}\text{number of eligible} \\ \text{responding}\end{array}\right] + \left[\begin{array}{c}\text{number of eligible} \\ \text{non-responding}\end{array}\right] + \left[\begin{array}{c}\text{estimated number of eligibles} \\ \text{from the unknowns}\end{array}\right]} \times 100$$

The justification for using this calculation method is that a proportion of the unknowns is likely to have been eligible if contact could have been made. This proportion of the unknowns is therefore treated as eligible non-respondents.

The estimated number of unknown eligibles is calculated as follows:

$$\left[\begin{array}{c}\text{Estimated number} \\ \text{of eligibles from the} \\ \text{unknowns}\end{array}\right] = \left[\begin{array}{c}\text{number} \\ \text{of} \\ \text{unknowns}\end{array}\right] \times \frac{\left[\begin{array}{c}\text{number of eligible} \\ \text{responding}\end{array}\right] + \left[\begin{array}{c}\text{number of eligible} \\ \text{non-responding}\end{array}\right]}{\left[\begin{array}{c}\text{number of eligible} \\ \text{responding}\end{array}\right] + \left[\begin{array}{c}\text{number of eligible} \\ \text{non-responding}\end{array}\right] + \left[\begin{array}{c}\text{number of} \\ \text{ineligibles}\end{array}\right]}$$

The response rate for children is calculated using the same approach as for adults, but where 'eligible' means the household contained at least one child and the definition of responding is that a child interview was conducted.

Coverage rate

The coverage rate is an alternative measure related to survey response and shows the extent to which a population has been involved in a survey. It provides information on the discrepancy between the sample (weighted by selection weight) and the population. It encompasses the impact of non-response rates, but also incorporates other factors such as being excluded or missed from the sample frame. For example, dwellings that have just been built may not be included in the sample frame, thereby contributing to under-coverage. The coverage rate is defined as the ratio of the sum of the selection weights for the survey to the known external population size.

Unlike the response rate, the coverage rate can be calculated without making any assumption about how many households with unknown eligibility were in fact eligible. Moreover, the coverage rate can usually be broken down in more detail than the response rate. However, definitional or operational differences between the survey scope and the external population size will affect the coverage rate (eg, differing definitions of usual residence). As a result, the response rate is generally used as the primary measure of the survey's quality. Some information on the coverage rate is included here to provide more detail on response, particularly response by ethnicity and age group.

Coverage rates also represent the factor by which the calibrated weighting process adjusts the initial selection weights in order to force agreement with benchmark data.

For details of the coverage rate in 2013/14, see Section 8.

Section 6: Weighting

Weighting of survey data ensures the estimates calculated from these data are representative of the target population. Most national surveys have complex sample designs whereby different groups have different chances of being selected in the survey. These complex designs are used for a variety of purposes, in particular to:

- reduce interviewer travel costs by ensuring the sample is geographically clustered, or 'clumped'
- ensure all regions of interest, including small regions, have a sufficient sample to enable adequate estimates
- ensure all sub-populations, in particular the Māori, Pacific and Asian populations, have a sufficient sample to enable adequate estimates.

To ensure no group is under- or over-represented in estimates from a survey, a method of calculating estimates that reflects the sample design must be used. Estimation weights are used to achieve this aim.

A weight is calculated for every respondent, and these weights are used in calculating estimates of population totals (counts), averages and proportions. Typically, members of groups that have a lower chance of selection are assigned a higher weight, so that these groups are not under-represented in estimates. Conversely, groups with a higher chance of selection receive lower weights. Also, groups that have a lower response rate (eg, young men) are usually assigned a higher weight so that these groups are correctly represented in all estimates from the survey.

The NZHS uses the calibrated weighting method to:

- reflect the probabilities of selection of each respondent
- make use of external population benchmarks (typically obtained from a population census) to correct for any discrepancies between the sample and the population benchmarks; this improves the precision of estimates and reduces bias due to non-response.

The NZHS data set is weighted every quarter separately from the population benchmarks. This means that each quarter can be used to produce valid population estimates.

Calculation of selection weights

The first step in producing calibrated weights is to calculate a selection probability (selection weight) for each respondent. It is crucial to calculate selection weights correctly to avoid bias in the final calibrated estimators.

Selection weights for the area-based sample and the electoral roll sample are calculated in different ways.

Area-based sample

- The probability of a meshblock i being selected in the area-based sample (A) is written as π_{Ai}. The values of π_{Ai} are greater than 0 for all meshblocks in the survey population.

- The probability of a dwelling being selected from a selected meshblock i in the area sample is $1/k_{Ai}$, where k_{Ai} is a skip assigned to each meshblock on the frame.

- The probability of an adult being selected from a selected dwelling j in a selected meshblock i is then $1/N_{ij(\text{adult})}$, where $N_{ij(\text{adult})}$ is the number of adults in the dwelling. Similarly, the probability of any particular child (if any in the household) being selected is $1/N_{ij(\text{child})}$, where $N_{ij(\text{child})}$ is the number of children in the dwelling.

Electoral roll sample

- The probability of a meshblock i being selected in the electoral roll sample (R) is written as π_{Ri}. The values of π_{Ri} are 0 for some meshblocks (those with few people who registered Māori descent on the electoral roll snapshot used in the sample design for that year).

- Dwellings are eligible for selection in the electoral roll sample if they have at least one adult registered with Māori descent in the electoral roll snapshot extracted for the enumeration quarter. ($E_{ij} = 1$ if meshblock i has $\pi_{Ri} > 0$ and dwelling j in this meshblock is eligible; $E_{ij} = 0$ otherwise.)

- A skip k_{Ri} is assigned to each meshblock and applied to eligible dwellings. The probability of an eligible dwelling being selected from meshblock i in the area sample is $1/k_{Ri}$, where k_{Ri} is a skip assigned to each meshblock on the frame.

- The probability of any particular adult being selected in the electoral roll sample from a selected dwelling j in a selected meshblock i is then $1/N_{ij(\text{adult})}$, and the probability of any particular child (if any in the household) being selected is $1/N_{ij(\text{child})}$.

Combined sample

The electoral roll sample and the area-based sample are selected according to the probabilities calculated using the above methods. The two samples of meshblocks do not overlap. The complete NZHS sample is defined as the union of the two samples. The probability of selection for any adult in dwelling j in meshblock i in the combined sample is therefore:

(1) $$\pi_{ij(\text{adult})} = E_{ij}\pi_{Ri}k_{Ri}^{-1}N_{ij(\text{adult})}^{-1} + \pi_{Ai}k_{Ai}^{-1}N_{ij(\text{adult})}^{-1} = \left(E_{ij}\pi_{Ri}k_{Ri}^{-1} + \pi_{Ai}k_{Ai}^{-1}\right)N_{ij(\text{adult})}^{-1}$$

Similarly, the probability of selection for any child in dwelling j in meshblock i in the combined sample is:

(2) $$\pi_{ij(\text{child})} = E_{ij}\pi_{Ri}k_{Ri}^{-1}N_{ij(\text{child})}^{-1} + \pi_{Ai}k_{Ai}^{-1}N_{ij(\text{child})}^{-1} = \left(E_{ij}\pi_{Ri}k_{Ri}^{-1} + \pi_{Ai}k_{Ai}^{-1}\right)N_{ij(\text{child})}^{-1}$$

The selection weights for adults and children are given by the reciprocal (inverse) of the above:

(3) $$d_{ij(\text{adult})} = \pi_{ij(\text{adult})}^{-1} = \left(E_{ij}\pi_{Ri}k_{Ri}^{-1} + \pi_{Ai}k_{Ai}^{-1}\right)^{-1}N_{ij(\text{adult})}$$

(4) $$d_{ij(\text{child})} = \pi_{ij(\text{child})}^{-1} = \left(E_{ij}\pi_{Ri}k_{Ri}^{-1} + \pi_{Ai}k_{Ai}^{-1}\right)^{-1}N_{ij(\text{child})}$$

Calibration of selection weights

Calibrated weights are calculated by combining the selection weights and population benchmark information obtained externally from the survey. The NZHS uses counts from Statistics New Zealand's estimated resident population for each calendar quarter, broken down by age, sex, ethnicity and socioeconomic position, as its benchmark population.

Calibrated weights are calculated to achieve two requirements.

1. The weights should be close to the inverse of the probability of selection of each respondent.

2. The weights are calibrated to the known population counts for a range of sub-populations (eg, age-by-sex-by-ethnicity categories). This means that the sum of the weights for respondents in the sub-population must exactly equal the known benchmark for the sub-population size.

The weights are chosen to minimise a measure of the distance between the weights and the inverse selection probabilities, provided that requirement 2 above is satisfied. Requirement 1 ensures that estimates have low bias, while requirement 2 improves the precision of estimates and achieves consistency between the survey estimates and external benchmark information.

A number of distance measures are in common use. A chi-square distance function (case 1 in Deville and Särndal 1992: 378) is used for the weighting of the NZHS, which corresponds to generalised regression estimation (also known as GREG). This distance function is slightly modified to force weights to lie within certain bounds, with the aim of avoiding extreme weights.

The inverse selection probability is sometimes called the initial weight. The final, calibrated weights are sometimes expressed as: final weight = initial weight * g-weight. The 'g-weight' indicates the factor by which calibration has changed the initial weight.

Weights for measurement participants

An extra set of statistical weights is calculated for the subset of participants who have their height and weight measured. Creating these weights follows exactly the same process as for the full sample. This consistent approach ensures that any bias due to lower participation in the measurement phase of the survey for particular demographic subgroups (such as age groups or ethnic groups) is accounted for in the final estimates for the survey. Analysis that uses the measurement data should always use this second set of weights.

Benchmark populations

The following benchmarks are used in the NZHS weighting:

- age (0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–74, 75+ years) by sex (male, female) for all people
- age (0–4, 5–9, 10–14, 15–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–64, 65+ years) by sex (male, female) for all Māori people
- total population by Pacific and non-Pacific (for adults)
- total population by Asian and non-Asian (for adults)
- total population by New Zealand Deprivation Index (NZDep2013) quintile.

Age, sex, ethnicity (Māori, Pacific, Asian, using self-identified total ethnicity) and socioeconomic position (NZDep2013) are included because these variables are related to many health conditions, are related to non-response, and are a key output classification for the survey. Current (within 12 months) population estimates at this level are available from Statistics New Zealand.

Quarterly calibration means that benchmarks are less detailed than would be possible if annual data sets were weighted. In particular, broader age groups are used for the Māori population benchmarks.

Benchmarks for the total Pacific and Asian populations

Benchmarks for the total Pacific and Asian populations are derived from Statistics New Zealand's Household Labour Force Survey. This large national survey (15,000 households surveyed per quarter) achieves a very high response rate (close to 90 percent). From this survey, Statistics New Zealand publishes quarterly estimates of the working-age Pacific and Asian populations.

Benchmarks for the NZDep2013 quintiles

Benchmarks for the quintiles of the New Zealand Index of Deprivation 2013 (NZDep2013) are derived by dividing the latest total population figures (of all ages, based on the 2013 Census data) into five groups of equal size. Refer to page 18 of this report for more details about the Index of Deprivation. Prior to the current 2013/14 NZHS, the New Zealand Index of Deprivation 2006 (NZDep2006, based on 2006 Census data) was used.

Calibration software and bounding of weights

The GREGWT SAS macro, produced by the Australian Bureau of Statistics, is used to calculate calibrated weights. The initial weights are the selection weights, re-scaled to sum to the population total. Final weights are constrained to be less than or equal to the smaller of 1625 and 2.5 times the initial weight. The constraint that weights must be less than 2.5 times the initial weight is equivalent to forcing the g-weight to be less than 2.5.

Section 7: Data processing and analysis

This section outlines the processes used to collect, check and output the data for the continuous NZHS.

Capture and coding

Questionnaire responses are entered directly on interviewers' laptops using Survey System computer-assisted personal interview (CAPI) software. Most of the questions have single-response options. However, a number of questions allow for multiple responses or require discrete numerical responses, such as age at the time of a specific event, or the number of visits to a specific medical professional. For these questions all responses are retained, with each response shown as a separate variable on the data file.

In addition, a number of questions in the questionnaire offer an 'other' category, where respondents can specify non-standard responses. Each 'other' category response is recorded (in free text). For each of these responses, the coders then choose one of the following options: re-categorise it to an existing code; code it to a newly set-up 'standard' code; or code it as 'other'. This coding is checked by both CBG and the Ministry of Health.

Security of information

Any information collected in the survey that could be used to identify individuals is treated as strictly confidential. Data are transferred daily from interviewers' laptops to CBG by a secure internet upload facility. The Ministry accesses the data through the CBG website using a secure log-in username and password.

The names and addresses of people and households that participate in the survey are not stored with response data. Unit record data are stored in a secure area and are only accessible on a restricted basis.

Checking and editing

CBG and the Ministry both undertake routine checking and editing of the data throughout the field period of the NZHS. In addition, the final unit record data sets provided to the Ministry are edited for range and logic. Any inconsistencies found are remedied by returning to the interviewer and, if necessary, to the respondent for clarification and correction.

Missing data due to non-response

Almost all questions have less than 1 percent missing data due to 'don't know' responses and refusals (the very few exceptions include questions on household composition and personal income). No explicit unit record or item imputation is used in the survey to deal with either unit record or item non-response.

Unit non-response is adjusted for in the calculation of weights, to the extent that this is possible using the weighting variables available (age, sex, ethnicity and NZDep).

Creation of derived variables

A number of derived variables are created on the NZHS data set. Standard definitions are used where possible. All derivations are thoroughly checked.

Derived variables such as educational qualification, labour-force status, body mass index and Short Form 36 (SF-36) score are based on commonly used or standard definitions. Other derived variables – such as a summary indicator of physical activity level that incorporates information on the intensity, duration and frequency of physical activity – are developed specifically for the analysis of the survey.

For ethnic group analyses, non-response is included as European/Other, as is 'New Zealander'.

More information on the derived variables in the NZHS will become available with the microdata release planned for June 2015.

Analysis methods

The data are analysed according to the following techniques for NZHS reports. References to weights here mean the calibrated weights discussed on pages 14–15.

Calculation of proportions

The proportion of the population who belong to a particular group (eg, the proportion of the population who have diabetes) is estimated by calculating the sum of the weights for the respondents in the group divided by the sum of the weights of all respondents.

The proportion of people in a population group who belong to a subgroup (eg, the proportion of Māori who have diabetes) is estimated by calculating the sum of the weights for the respondents in the subgroup (Māori who have diabetes) divided by the sum of the weights for the respondents in the population group (Māori).

In each case, respondents with missing data (eg, who either refused to say, or didn't know, whether they had diabetes) are excluded from both the numerator and the denominator of the calculation.

Calculation of totals (counts)

Estimates of totals are given by calculating the sum, over all the respondents, of the weight multiplied by the variable of interest. For example, the estimate of total number of people with diabetes in the whole population would be given by the sum, over all respondents, of the weight multiplied by a binary variable indicating which respondents have diabetes. This is equivalent to the sum of the weights for the respondents who have diabetes in the population.

To account for item non-response, an additional factor is calculated using the sum of the weighted number of respondents to the variable of interest and the weighted number of item non-respondents, divided by the weighted number of respondents.

Calculation of averages (means)

Estimates of population averages, such as the average number of visits to a general practitioner (GP), are determined by calculating the sum, over all respondents, of the weight multiplied by the variable of interest, divided by the sum of the weights.

Sometimes the average within a group is of interest; for example, the average number of visits to a GP by males. The estimate is given by calculating the sum, over respondents in the group, of the weight multiplied by the variable of interest divided by the sum of the weights of respondents in the group.

In each case, respondents with missing data (eg, who either refused to say or didn't know how many GP visits they had made) are excluded from both the numerator and the denominator of the calculation.

Ethnicity

The reports use *total response ethnicity* to define ethnic groups. Total response ethnicity classifies a person in all the ethnic groups they identify with. This means that people can appear in more than one ethnic group.

Neighbourhood deprivation

Neighbourhood deprivation refers to the New Zealand Index of Deprivation 2013 (NZDep2013), which measures the level of socioeconomic deprivation for each neighbourhood (meshblock) according to a combination of the following 2013 Census variables: income, benefit receipt, transport (access to car), household crowding, home ownership, employment status, qualifications, support (sole-parent families) and access to a telephone (Atkinson et al 2014).

Survey data are generally presented for NZDep2013 quintiles 1–5. Quintile 1 represents the 20 percent of small areas with the lowest levels of deprivation (the least deprived areas) and quintile 5 represents the 20 percent of small areas with the highest level of deprivation (the most deprived areas).

To explore the association of selected indicators with neighbourhood deprivation, the *relative index of inequality* is used. This index is calculated by first using data from all quintiles to calculate a line of best fit (regression line), adjusted for age group, sex and ethnic group. The points on the regression line corresponding to the most and least deprived areas are used to calculate the relative index of inequality.

Age standardisation

Crude (unadjusted) rates for estimates of the prevalence in the total population and by age group are presented in the reports. However, age is an important determinant of health, so populations with different age structures (such as men and women, due to women's longer life expectancy) may have different rates due to these age differences.

Age standardisation is performed by the direct method using the World Health Organization (WHO) world population age distribution (Ahmad et al 2000). This statistical method of standardising for age is used in analyses by sex, ethnic group and neighbourhood deprivation (NZDep2013), and for comparisons between the different health surveys. Results for children are age standardised to the population younger than 15 years, and results for adults are age standardised to the population aged 15 years and over.

Rate ratios

The reports present adjusted rate ratios for the following comparisons:

- men and women
- Māori and non-Māori (for total, men, women)
- Pacific and non-Pacific (for total, men, women)
- Asian and non-Asian (for total, men, women)
- people living in the most and least deprived areas (for total, men, women).

The rate ratios can be interpreted in the following ways.

- A value of 1.00 shows that there is no difference between the group of interest (eg, men) and the reference group (eg, women).
- A value higher than 1.00 shows that the result is higher for the group of interest than for the reference group.
- A value lower than 1.00 shows that the result is lower for the group of interest than for the reference group.

In the neighbourhood deprivation comparisons, the rate ratio refers to the relative index of inequality.

Depending on the comparison, the rate ratio is adjusted appropriately for other demographic factors such as age, sex and ethnic group that may be influencing (confounding) the comparison.

- The sex comparison is adjusted for age.
- The ethnic comparisons are adjusted for age and sex.
- The deprivation comparisons are adjusted for age, sex and ethnic group.

The adjusted rate ratio reflects the true comparison between the group of interest (eg, men) and the reference group (eg, women) by controlling for the impact of the other factors.

Confidence intervals

Ninety-five percent confidence intervals are used to represent the sample error for estimates. A 95 percent confidence interval means there is a 95 percent chance that, if we were to survey the whole population, the true value would lie between the lower and upper confidence interval values.

Some analysts assess whether two estimates differ significantly by seeing whether their confidence intervals overlap or not. This procedure is known to be overly conservative, resulting in a substantial degrading of statistical power, with some significant differences incorrectly assessed as insignificant. If the confidence intervals do not overlap, then it can be concluded that the estimates differ significantly. However, when they do overlap, it is still possible that there is a significant difference. In this case, a t-test is used to correctly test the statistical significance of differences between NZHS estimates.

Percentiles

To calculate variances (and hence confidence intervals) using replicate weights for percentiles (including medians), the Woodruff method is used (Woodruff 1952).

Time trends

Where possible, the results of indicators presented in the current report are compared with the corresponding results of the 2012/13 NZHS and the 2006/07 NZHS to examine whether an indicator shows a significant increase or decrease. This is referred to as 'time trends' in the annual report. The time trend analyses are based on age-standardised p-values.

Small numbers

Suppression of estimates

Small samples can affect both the reliability and the confidentiality of results. Problems with reliability arise when the sample becomes too small to adequately represent the population from which it has been drawn. Problems with confidentiality can arise when it becomes possible to identify an individual, usually someone in a subgroup of the population within a small geographical area.

To ensure the survey data presented are reliable and the confidentiality of the participants is protected, data have only been presented when there are at least 30 people in the denominator (the population group being analysed). Care has been taken to ensure that no participant can be identified in the results.

Calculation of confidence intervals

In many cases, confidence intervals based on standard normal approximation do not work well when estimating small proportions. When the sample size for the sub-population being estimated is small, the symmetrical behaviour of these confidence intervals can be unrealistic and can even lead to confidence intervals containing negative values.

The Korn and Graubard (1998) method is used to calculate confidence intervals in any of the following circumstances:

- the numerator (number of respondents with the variable of interest) is less than 30
- the lower confidence interval results in a value less than 0
- the upper confidence interval results in a value greater than 100.

In these circumstances, the resulting confidence intervals can and should be asymmetrical.

Section 8: New Zealand Health Survey 2013/14

This section provides some field-related data specific to the data collection and analysis of the NZHS in 2013/14. Appendix 1 contains some information on the recent health surveys carried out in 2012/13 and 2006/07.

2013/14 NZHS module

Table 3 outlines the NZHS module topics for 2013/14.

Table 3: NZHS module topics, 2013/14

Adult module topics	Child module topics
Long-term conditions	Long-term conditions
Health status	Health status
Disability status	Disability status
Living standards	Living standards
Housing quality	Housing quality

Further details on the questionnaires for 2013/14 can be found in the *Content Guide 2013/14: New Zealand Health Survey* (Ministry of Health 2014b). The survey questionnaires can be found with the *Annual Update of Key Results 2013/14: New Zealand Health Survey*.

Data collection

In the second year of the continuous NZHS, 1 July 2013 to 30 June 2014, a total of 13,309 adults and 4699 children took part in the survey. Table 4 shows the number of participants selected each quarter in 2013/14.

Table 4: Number of survey participants, by quarter, 2013/14

	Adults		Children	
	Number	**Percentage of total participants**	**Number**	**Percentage of total participants**
Quarter 1 (July–September 2013)	3505	26	1294	27
Quarter 2 (October–December 2013)	3341	25	1133	24
Quarter 3 (January–March 2014)	3300	25	1166	25
Quarter 4 (April–June 2014)	3163	24	1106	24
Total (July 2013–June 2014)	13,309	100	4699	100

Call pattern

As previously described in Section 4, the call pattern used in the NZHS is an important part of achieving a high response rate. In 2013/14, surveyors followed a proven call approach, including visiting meshblocks at different times and on different days depending on the area they were working in. For about 90 percent of households, the first (or only) interview took place within seven calls (Figure 1).

Figure 1: Proportion of households agreeing to first interview, by number of calls, 2013/14
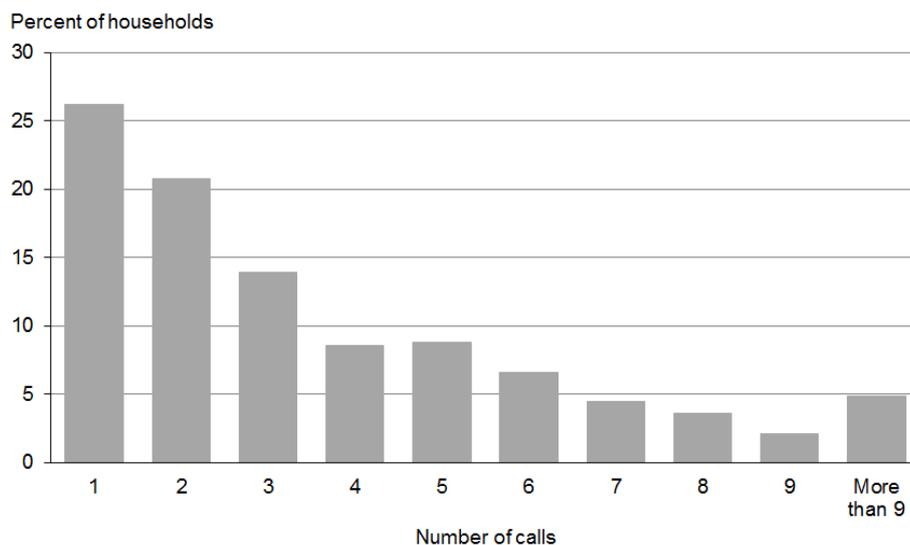
Response rates

The current NZHS is well received by the public, with the weighted response rates in 2013/14 being 80 percent for adults and 85 percent for children. The corresponding rates in 2012/13 were the same for both adults and children (80 percent for adults and 85 percent for children).

The collection of measurement data (including height and weight) had a slightly lower response rate among survey respondents, and therefore the measurement data were reweighted so that they represented the total population.

Coverage rates

In 2013/14 the coverage rate was 54 percent for adults and 63 percent for children. The corresponding rates in 2012/13 were 59 percent for adults and 69 percent for children. The rates in 2011/12 were 54 percent for adults and 68 percent for children.

Table 5 shows the coverage rates in 2013/14, by ethnic group and neighbourhood deprivation. There were good coverage rates for Māori and Pacific peoples (similar to the rate for all adults). Coverage rates were high for children, reflecting high rates for adults in the typical parenting age range (Figure 2).

Table 5: Coverage rates, children and adults combined, 2013/14

Population	Coverage rate (%)
Māori	65
Pacific	60
Asian	53
NZDep2013 Quintile 1	53
NZDep2013 Quintile 2	48
NZDep2013 Quintile 3	70
NZDep2013 Quintile 4	58
NZDep2013 Quintile 5	52

Figure 2 shows the coverage rates by age and sex. The pattern for Māori in Figure 3 is similar to the overall pattern.

**Figure 2: Coverage rates (%) for total population, by age group and sex, 2013/14**
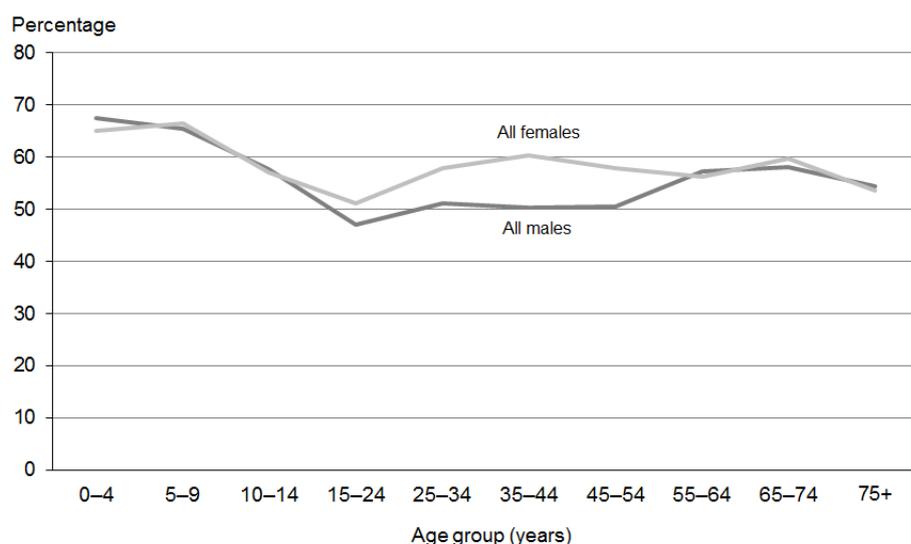
Figure 3: Coverage rates (%) for Māori, by age group and sex, 2013/14
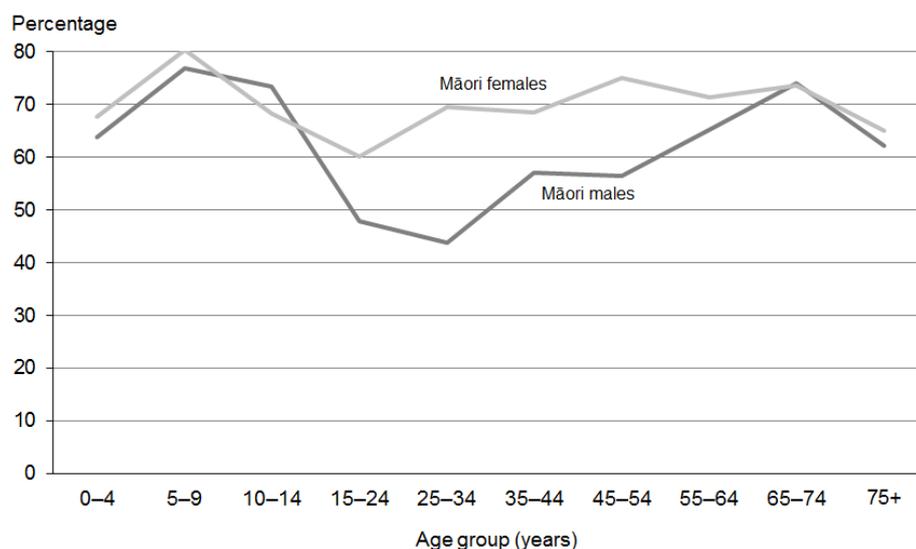
Final weights

Section 6 explained how the calibrated weights were calculated. Table 6 gives basic descriptive information on the final weights calculated for the 2013/14 survey.

The g-weights are the ratios of the final weights to the initial selection weights. The mean g-weight is approximately 1.8. This means the calibrated weights, which were calculated using population benchmark information, have changed the initial selection weight by an average factor of 1.8.

Table 6: Final weights, 2013/14

	Final weights
Minimum	5
Median	196
90th percentile	539
95th percentile	718
99th percentile	1175
Maximum	1625
Coefficient of variation (CV%)	87.0
Approximate design effect due to weighting (1 + CV^2)	1.76

Sample sizes

Tables 7 to 10 show the 2013/14 NZHS sample sizes and the total usually resident population counts, by sex, ethnicity, age and NZDep2013 quintile.

Table 7: Sample sizes and population counts for children and adults, by sex, 2013/14

Population group	Sex	Interviews	Measurements (2+ years)	Population count
Children (0–14 years)	Boys	2410	1828	466,485
	Girls	2289	1732	443,518
	Total	4699	3560	910,003
Adults (15 years and over)	Men	5789	5592	1,728,990
	Women	7520	7002	1,845,857
	Total	13,309	12,594	3,574,847

Table 8: Sample sizes and population counts for children and adults, by ethnic group, 2013/14

Ethnic group (total response)	Population group	Interviews	Measurements (2+ years)	Population count
European/Other	Children	3101	2376	637,674
	Adults	9897	9411	2,725,269
Māori	Children	1779	1341	237,811
	Adults	2837	2648	449,842
Pacific	Children	677	501	117,309
	Adults	802	747	195,000
Asian	Children	522	384	113,475
	Adults	1089	1022	441,750

Table 9: Sample sizes and population counts, by age group, 2013/14

Age group (years)	Interviews	Measurements (2+ years)	Population count
0–4	1787	872	309,837
5–9	1479	1378	304,133
10–14	1433	1310	296,033
15–24	1619	1518	635,025
25–34	2112	1924	561,833
35–44	2326	2222	588,637
45–54	2271	2182	627,588
55–64	2049	1967	520,288
65–74	1666	1598	370,340
75 and over	1266	1183	271,138

Table 10: Sample sizes and population counts, by NZDep2013 quintile, 2013/14

NZDep2013 quintile	Population group	Interviews	Measurements (2+ years)	Population count
Quintile 1 (least deprived neighbourhoods)	Children	535	430	172,716
	Adults	1631	1550	724,254
Quintile 2	Children	594	461	164,582
	Adults	1917	1822	732,388
Quintile 3	Children	1090	806	160,333
	Adults	3350	3186	736,637
Quintile 4	Children	1133	849	186,040
	Adults	3174	3011	710,930
Quintile 5 (most deprived neighbourhoods)	Children	1347	1014	226,332
	Adults	3237	3025	670,638

References

Ahmad O, Boschi-Pinto C, Lopez A, et al. 2000. *Age-standardization of rates: A new WHO standard*. Geneva: World Health Organization.

Atkinson J, Salmond C, Crampton P. 2014. *NZDep2013 Index of Deprivation*. Wellington: Department of Public Health, University of Otago.

Clark RG, Templeton R, McNicholas A. 2013. Developing the design of a continuous national health survey for New Zealand. *Population Health Metrics* 11(1): 25. URL: www.pophealthmetrics.com/content/11/1/25

Deville JC, Särndal CE. 1992. Calibration estimators in survey sampling. *Journal of the American Statistical Association* 87: 376–82.

Korn EL, Graubard BI. 1998. Confidence intervals for proportions with small expected number of positive counts estimated from survey data. *Survey Methodology* 24(2): 193–201.

Ministry of Health. 2008. *A Portrait of Health: Key results of the 2006/07 New Zealand Health Survey*. Wellington: Ministry of Health.

Ministry of Health. 2011. *The New Zealand Health Survey: Sample design, years 1–3 (2011–2013)*. Wellington: Ministry of Health. URL: www.health.govt.nz/publication/new-zealand-health-survey-sample-design-years-1-3-2011-2013

Ministry of Health. 2012. *New Zealand Health Survey Methodology Report 2012*. Wellington: Ministry of Health.

Ministry of Health. 2013. *New Zealand Health Survey Methodology Report 2012/13*. Wellington: Ministry of Health.

Ministry of Health. 2014a. *Annual Update of Key Results 2013/14: New Zealand Health Survey*. Wellington: Ministry of Health.

Ministry of Health. 2014b. *Content Guide 2013/14: New Zealand Health Survey*. Wellington: Ministry of Health.

Statistics New Zealand. 1998. *Protocols of Official Statistics*. Wellington: Statistics New Zealand.

Statistics New Zealand. 2007. *Principles and Protocols for Producers of Tier 1 Statistics*. Wellington: Statistics New Zealand.

Woodruff RS. 1952. Confidence intervals for medians and other position measures. *Journal of the American Statistical Association* 47: 635–46.

Appendix 1: Previous New Zealand Health Surveys

To determine any changes in the prevalence of indicators over time, the *Annual update of key results 2013/14: New Zealand Health Survey* shows results comparing the current NZHS with the previous surveys conducted in 2006/07 and 2012/13. This section gives a brief description of those two surveys.

2006/07 New Zealand Health Survey

The target population for the 2006/07 NZHS was the usually resident civilian population of all ages living in permanent private dwellings in New Zealand. An area-based frame of Statistics New Zealand meshblocks was used as the sample frame. Māori, Pacific and Asian peoples were oversampled.

Data were collected from October 2006 to the end of November 2007 using computer-assisted, face-to-face interviewing. The total response rate for the survey was 68 percent for adults and 71 percent for children. A total of 12,488 adults and 4921 children took part in the survey. The survey included 11,632 European/Other people, 5143 Māori, 1831 Pacific people and 2255 Asian people of all ages.

For full details on the methodology of the 2006/07 NZHS, see *A Portrait of Health: Key results of the 2006/07 New Zealand Health Survey* (Ministry of Health 2008).

2012/13 New Zealand Health Survey

This is the previous NZHS, which was conducted in 2012/13. The target population was the New Zealand usually resident population of all ages, including those living in non-private accommodation to improve the coverage of older people.

The sample design was the same as for the 2013/14 NZHS, which involves a multi-stage, stratified, probability-proportional-to-size sampling design. The sample was selected from a dual sampling frame, whereby the participants were selected from an area-based sample and a list-based electoral roll sample.

Data were collected from July 2012 to July 2013 using computer-assisted, face-to-face interviewing. The total response rate for the survey was 80 percent for adults and 85 percent for children. A total of 13,009 adults and 4485 children took part in the survey. The survey included 12,598 European/Other people, 4230 Māori, 1411 Pacific people and 1469 Asian people of all ages.

For full details on the methodology of the 2012/13 NZHS, see the *New Zealand Health Survey Methodology Report 2012/13* (Ministry of Health 2013).