Methodology Report 2014/15

New Zealand Health Survey

New Zealand Health Survey

Authors

This report was written by Deepa Weerasekera, Steven Johnston, Bridget Murphy (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

MINISTRY OF
HEALTH
MANATŪ HAUORA

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 eight Tier 1 statistics from the NZHS are smoking (current), past-year (alcohol) drinking, hazardous (alcohol) drinking, obesity, unmet need for GP due to cost, unfilled prescription due to cost, self-rated health and mental health status (psychological distress).

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 4 of the continuous NZHS (data collected between July 2014 and June 2015) is included in Section 9 of this report. The corresponding information for years 1, 2 and 3 of the NZHS can be found in the previous methodology reports.[1]

---

[1] See www.health.govt.nz/nz-health-statistics/national-collections-and-surveys/surveys/current-recent-surveys/new-zealand-health-survey

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 2014/15* contains further information on these objectives (Ministry of Health 2015b).

The NZHS has been carefully designed to minimise the impact on respondents. Features for this purpose 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 New Zealand Health and Disability Multi-Region Ethics Committee granted approval for the NZHS (MEC/10/10/103) in 2011.

Section 2: Survey content

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 0–14 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 2014/15* (Ministry of Health 2015b).

Core component

Most of the core questions for both adults and children are 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, 2011/12–2015/16

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 Biomedical tests	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.

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 proportion of the target population is 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 facilities)
- households in remote areas, including areas (meshblocks) with fewer than six occupied dwellings and those located off the main islands of New Zealand.

Included in the survey population are people living in aged-care facilities (rest homes) and those temporarily living away from their household in student accommodation (university hostels and boarding schools).

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 an overview of the sample design, see Clark et al (2013). 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).

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 peoples 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 the frame occupied dwellings count divided 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.

Any aged-care facilities in the selected meshblocks are included in the area-based sample by first dividing them into 'accommodation units', typically consisting of an individual or couple living together in the institution. Accommodation units are then treated as households in the sampling process, although at most five accommodation units are selected from a single facility.

Students living away from home in university hostels and boarding schools are eligible to be selected via their family's house, if they still consider this to be their home. If selected, arrangements are made to survey them either when they are next at home or at their student accommodation.

Electoral roll sample

The electoral roll is another sampling 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 containing at least one person who has self-identified as having Māori ancestry. The sample of meshblocks is selected so that it does not overlap with the sample from the area-based sample. 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). 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.

The 2014/15 adult module on Sexual and Reproductive Health was self-completed by the respondent. If the respondent was unable to use the laptop due to IT literacy or language issues, the module was skipped.

Pilot study

A pilot study was carried out with 100 respondents from 11 meshblocks in Auckland, Tauranga and Taranaki prior to the main data collection for the 2014/15 NZHS. The purpose of the pilot was to determine average survey timings for different groups and refine the survey instruments, materials, operations and processes. No substantive changes were made following the pilot study.

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 for the meshblock.

Invitation to participate

The NZHS is voluntary, relying on the goodwill of participants, and consent is obtained without coercion or inducement. CBG uses the NZ Post address database to post each selected 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 from each selected household 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 requires the respondent to confirm that they have read and understood the information pamphlet and that they know they can ask questions at any time and can contact CBG or the Ministry of Health for further information. The consent form also states that:

- the respondent can stop the interview at any time and they don't have to answer every question
- their participation is confidential and no identifiable information will be used in any reports
- their answers are protected by the Privacy Act 1993.

In 2014/15, adults aged 16–74 years were also provided with a separate information sheet relating to the Sexual and Reproductive Health module. This was presented before the consent form was signed and contained additional information relevant to the module.

The consent form also includes a request for an interpreter, if required (in any of a range of different languages). The respondent may elect for a friend or family member (aged 15 years or over) to act as the interpreter during the survey, or else CBG arranges for a professional interpreter to be provided. Attempts are also made to match respondents and interviewers by ethnicity and sex when requested.

Where a selected adult respondent is unable to provide consent themselves, a legal guardian is permitted to consent and complete the survey on their behalf if the legal guardian agrees to participate.

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 thank you 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. Adult respondents who completed the Sexual and Reproductive Health module were also given a separate thank you card at the conclusion of the interview, listing support organisations relevant to the module content.

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 or 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. In addition, adults aged 15 years and over are 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 the interviewer holds against the corner of a wall or door. The headboard is lowered until it reaches the participant's head, at which point the laser is activated to take a measurement. The laser design was trialled and refined in early 2012 before it was used for the entire survey from 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 measurements, then a third measurement is taken for accuracy. To align with international standards, the final height, weight and waist measurements used for analysis are calculated for each respondent by taking the mean of the two closest 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.

Several techniques are used to ensure the quality of the objective measurement equipment. Surveyors immediately report faulty equipment to CBG management and replacements are sent immediately. Equipment is also checked by CBG management at least twice every year in the field.

Additionally the equipment is checked and re-calibrated at the time of the annual module change training, where:

- the electronic weighting scales and automatic blood pressure machines are re-calibrated by a manufacturer-approved agent
- the lasers are checked against a known fixed height to ensure that they are measuring correctly, and that they are still programmed to the correct settings.

Surveyors are retrained annually and must pass a re-certification assessment to ensure that the required skill levels are maintained.

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 2014/15 the final weighted response rates were 79 percent for adults and 83 percent for children.
>
> For more details on the response rates for 2014/15, see Section 9.

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 NZHS calculates a weighted response rate. The weight of each household reflects the probability of the household being selected into the sample, so that the weighted response rate describes the success of the survey in terms of achieving cooperation from the population being measured.

For adults, the response rate calculation classifies all selected households into the following four groups:

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 provided 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-responding.

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 '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 responding 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, in this way contributing to under-coverage.

The coverage rate is defined as the ratio of the sum of the selection weights for the survey respondents 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, including by individual characteristics. However, definitional or operational differences between the survey scope and the external population size (eg, differing definitions of usual residence) will affect the coverage rate. 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 selection weights in order to force agreement with the calibration benchmarks (see Section 7 for more on calibration).

For details of the coverage rates in 2014/15, see Section 9.

Section 6: Data processing

Capture and coding

Questionnaire responses are entered directly on interviewers' laptops using The Survey System computer-assisted personal interview (CAPI) software.

Most questions have single-response options 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. However, a number of questions allow for multiple responses. 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.

Ethnicity is self-defined, and respondents are able to report affiliation with more than one ethnic group using the Statistics New Zealand standard ethnicity question. Responses to the ethnicity question are coded to level 3 of the 2005 standard ethnicity classification.

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

Unit non-response is adjusted for in the calculation of weights, as described in Section 7. Weighting is also used to adjust for non-response to the measurement phase of the interview.

Almost all questions have less than 1 percent of missing data due to 'don't know' responses and refusals. The exceptions are:

- personal income and household income (between 10% and 20% non-response)
- questions that ask for the cost of the respondent's last GP or nurse visit (around 3–4% non-response).

Where a respondent does not provide their date of birth or their age in years, age is imputed as the midpoint of the age group they have provided. No other imputation is used to deal with item non-response.

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 Body Mass Index (BMI), Alcohol Use Disorders Identification Test (AUDIT) and level of psychological distress (K10) 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.

Outliers

Respondents with height and weight measurements that lead to a calculated BMI of less than 10 or greater than 80 are treated as non-respondents to the measurement phase of the interview.

Respondents who report more than 112 hours of physical activity per week (an average of 16 hours per day) are excluded from the derived summary measure of physical activity.

Ethnicity

Ethnic group variables are derived using the concept of **total response ethnicity** (Statistics New Zealand 2005). This means that respondents can appear in, and contribute to the published statistics for, more than one ethnic group.

NZHS reports generally provide statistics for the following four ethnic groups: Māori, Pacific, Asian, and European/Other. The 'Other' ethnic group (comprising mainly Middle-Eastern, Latin-American and African ethnicities, and those who answer 'New Zealander') has been combined with European to avoid problems with small sample sizes.

Neighbourhood deprivation

Neighbourhood deprivation refers to the New Zealand Index of Deprivation 2013 (NZDep2013), developed by researchers at the University of Otago (Atkinson et al 2014). NZDep2013 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.

NZHS reports generally use NZDep2013 quintiles, where 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).

A small number of meshblocks do not have a value for NZDep2013. If any of these meshblocks are selected in the NZHS, the respondents are assigned to quintile 3 (ie, the middle quintile) for weighting and analysis purposes.

Section 7: 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
- ensure all regions of interest, including small regions, have a sufficient sample size for adequate estimates to be made
- ensure important sub-populations, in particular the Māori, Pacific and Asian ethnic groups, have a sufficient sample size for adequate estimates to be made.

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 based on the 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.

Data from each calendar quarter of the NZHS data set are weighted separately to population benchmarks for that quarter. This means that each quarter's data 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 (and hence 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 any particular 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 fewer than three households with residents 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 electoral roll 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:

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

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

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

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

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

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

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 calibrated weights are calculated in such a way as to minimise a measure of the distance between the calibrated 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) is used for calibrating the NZHS weights, 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.

Population benchmarks

The following population 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
- adult population by Pacific and non-Pacific
- adult population by Asian and non-Asian
- 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.

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 Māori population

Quarterly benchmarks for the Māori population are constructed for the NZHS by projecting forward the annual (mid-year) population estimates for Māori released by Statistics New Zealand.

Using the Māori population estimates and total population estimates as at 30 June, the proportion of the total population who are Māori is calculated for each five-year age by sex group. Then these proportions are applied to quarterly total population estimates, by age and sex, for the subsequent four quarters.

For example, the proportion of each age by sex group who are Māori as at 30 June 2011 is used to construct estimates of the Māori population by age and sex in each of the quarters ending 30 September 2011, 31 December 2011, 31 March 2012 and 30 June 2012.

Benchmarks for the Pacific and Asian populations

Quarterly benchmarks for the adult Pacific and Asian populations are derived from Statistics New Zealand's Household Labour Force Survey. This large national survey, of 15,000 households per quarter, achieves a very high response rate (close to 90 percent).

The Household Labour Force Survey publishes quarterly estimates of the working-age (aged 15 years and over) Pacific and Asian populations. From these estimates the proportions of the adult population who are Pacific and Asian are obtained for each quarter. Some of the quarter-to-quarter variation in these proportions is smoothed out by applying a moving average over the quarterly figures. The final smoothed proportions are applied to the total adult benchmark for the corresponding quarter to give quarterly benchmarks for Pacific and Asian adults.

Benchmarks for the NZDep2013 quintiles

Benchmarks for the quintiles of NZDep2013 are derived by dividing the latest total population figures (of all ages) into five groups of equal size.

The calibration for the 2011/12 and 2012/13 surveys used benchmarks for the New Zealand Index of Deprivation 2006 (NZDep2006) based on 2006 Census data, while the surveys from 2013/14 onwards have used NZDep2013.

Calibration software and bounding of weights

The GREGWT SAS macro, developed by the Australian Bureau of Statistics, is used to calculate the calibrated weights. The input weights are the selection weights, first re-scaled to sum to the overall population benchmark. Final weights are constrained to be less than or equal to the smaller of 1,625 and 2.5 times the input weight.

Jackknife replicate weights

The NZHS uses the delete-a-group jackknife method (Kott 2001) to calculate standard errors for survey estimates.

One hundred jackknife replicate weights are produced for every respondent in the survey, in addition to the final calibrated weight. Each replicate weight corresponds to removing a group of meshblocks from the sample and re-weighting the remaining sample using exactly the same approach as that used to construct the weights for the full sample, including calibration to the same population benchmarks.

For any weighted estimate calculated from the survey, 100 jackknife replicate estimates can also be calculated using the 100 jackknife weights. The standard error of the full sample estimate is based on the variation in the replicate estimates.

A number of statistical analysis packages, including SAS, Stata and R, can calculate standard errors using jackknife weights.

Weights for measurement data

An additional set of estimation weights (and corresponding jackknife weights) has been created specifically for analysis of the measurements collected from respondents as part of the core NZHS interview. Height and weight measurements are obtained from around 96 percent of eligible adult and 92 percent of eligible child respondents. As variables derived from height and weight are key outputs from the survey, it is useful to have this additional set of estimation weights to compensate for the additional non-response to these items.

The extra set of weights is calculated for the subset of respondents who have their height and weight measured. Creating these estimation 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.

These estimation weights are also used for analysis involving waist and blood pressure measurements. Waist and blood pressure measurements are obtained from almost all respondents who have had their height and weight measured.

Section 8: Analysis methods

Estimation of prevalences, totals and means

Most statistics published in NZHS reports are presented as prevalences, totals or means.

Prevalences are survey estimates of the proportion of people with a particular characteristic, such as a specific health condition, behaviour or outcome. **Totals** are survey estimates of the number of people with a particular characteristic. **Means** are survey estimates of the average per person of some numeric quantity.

A description of the calculation method for each of these types of statistics follows. References to weights here mean the final calibrated weights discussed in Section 7.

Adjusting for item non-response

Before calculating prevalences, totals or means for a particular variable, an adjustment is made to the final weights to account for respondents who answered 'don't know' to or refused to answer the question.

The adjustment rates up the final weights of the respondents who gave a response to the item, to represent the final weights of the respondents who answered don't know or refused to answer for the variable in question. This is carried out within cells defined by sex and age group (10 year age groups for adults and 5 year age groups for children), making use of some information on what type of respondents are more likely to be item non-respondents to this variable. Then the item non-respondents can be safely left out of the calculation of prevalences, totals or means for the variable.

The adjustment is most important for totals, so that item non-response does not lead to the under-estimation of the number of people having the particular condition or behaviour. The effect will usually be very small for prevalences and means – that is, prevalences and means using the adjusted weights will be very similar to those using the final calibrated weights.

The adjustment is done 'on the fly' in the sense that the item-specific weights are created and used in estimation, but are not kept on the survey data set.

Calculation of prevalences (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 for 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 the 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.

Calculation of means (averages)

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 for the 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.

Suppression of small sample sizes

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.

Comparisons between population groups

Age standardisation

NZHS reports mainly focus on presenting crude (unadjusted) rates for estimates of the prevalence in the total population and by age group (age-specific rates).

However, age is an important determinant of health, so population groups with different age structures (such as men and women, whose age structures differ due to women's longer life expectancy) may have different rates due to these age differences. This means that comparisons of crude rates over time and between groups may be misleading if the age structure differs in the groups being compared.

One approach to making more meaningful comparisons between groups is to compare age-specific rates. Alternatively it can be useful to summarise a set of age-specific rates for a group into a single age-independent measure. This is achieved by a process called **age standardisation**.

Age standardisation in NZHS reports is performed by **direct standardisation** using the World Health Organization (WHO) world population age distribution (Ahmad et al 2000). The direct method calculates an age-standardised rate, which is a weighted average of the age-specific rates, for each of the population groups to be compared. The weights applied represent the relative age distribution of the WHO population. This provides a single summary rate for each of the population groups being compared that reflects the rate that would have been expected if the group had had an age distribution identical to the WHO population.

The age-standardised prevalence rate (ASR) is given by:

$$ASR = \sum r_i \, (n_i / \sum n_i),$$

where n_i is the population in the ith age group of the standard population and r_i is the prevalence rate in the ith age group from the survey.

Age-standardised rates are provided in some tables to help make comparisons by sex, ethnic group and neighbourhood deprivation (NZDep2013), and comparisons between survey years.

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.

Adjusted rate ratios

NZHS reports also present comparisons between population groups as **rate ratios**; that is, as the ratio of the prevalence estimates of the two groups.

Rate ratios are used for the following comparisons:

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

In keeping with the use of total response ethnicity to present prevalences by ethnic group, ethnic comparisons are presented such that Māori are compared with non-Māori, Pacific with non-Pacific and Asian with non-Asian. For this purpose, all respondents who identified as Māori are included in the Māori group; all other respondents are included in the non-Māori group. Similar groups are formed for Pacific and Asian ethnic groups.

Rate ratios can be interpreted in the following way.

- A value of 1 shows that there is no difference in prevalence between the group of interest (eg, men) and the reference group (eg, women).
- A value higher than 1 shows that the prevalence is higher for the group of interest than for the reference group.
- A value lower than 1 shows that the prevalence is lower for the group of interest than for the reference group.

The rate ratios presented in NZHS reports are adjusted for differences in demographic factors between the groups being compared 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 comparison is adjusted for age, sex and ethnic group.

Adjusting for potential confounding factors makes comparisons more accurate and meaningful, as the adjustment removes the effect of these confounding factors.

Adjusted rate ratios are calculated using the **predictive margins** approach of Korn and Graubard (1999), which Bieler et al (2010) call **model-adjusted risk ratios**.

- A statistical model is fitted to the data – a logistic model for a binary outcome or a log-linear regression model when the outcome variable is a count. The variable defining the groups to be compared, and the adjustment variables, are explanatory variables in the model.

- The parameters of the model are used to estimate the prevalence (or mean) as if all the respondents belong to the group of interest (eg, were all male) but otherwise each keeps their own values for the adjustment variables in the model (eg, age). That is, the prevalence being estimated is for a population where the distributions of the adjustment variables match the distribution of the respondents, except that everyone belongs to the group of interest.

- Once the model-adjusted prevalences for the group of interest (eg, men) and the comparison group (eg, women) have been estimated in this way, their ratio can be calculated.

In the neighbourhood deprivation comparisons, the rate ratio refers to the **relative index of inequality** (Hayes and Barry 2002). This measure is used instead of simply comparing the most deprived quintile with the least deprived quintile. It 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 rate ratio that is presented in the reports. This method has the advantage of using data from all the NZDep2013 quintiles to give an overall test for trend (gradient) by neighbourhood deprivation, rather than only using the data from quintiles 1 and 5.

While total response ethnicity is used to report ethnic group statistics in the NZHS reports, a prioritised ethnicity variable is used when adjusting for ethnicity in the regression models. Using prioritised ethnicity in the models simplifies the modelling process and gives results similar to including total response ethnicity variables in the models. The priority ordering of ethnic groups used is: Māori, Pacific, Asian, European/Other.

Confidence intervals and statistical tests

Ninety-five percent confidence intervals are used in NZHS reports to represent the sampling error associated with the statistics; that is, the uncertainty due to selecting a sample to estimate values for the entire population. A 95 percent confidence interval for a statistic is constructed in such a way that, under a hypothetical scenario where selecting the sample could be repeated many times, 95 percent of the confidence intervals constructed in this way would contain the true population value.

Calculation of confidence intervals

In most cases, confidence intervals presented in NZHS reports are calculated using the usual normal approximation. The upper and lower limits of the 95 percent confidence interval are found by:

> estimate ± 1.96 x standard error of the estimate.

However, confidence intervals based on the normal approximation sometimes do not work well when estimating small proportions. In these cases, the symmetrical behaviour of these normal 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 more appropriate confidence intervals in any of the following circumstances:

- the prevalence estimate is less than 5 percent or greater than 95 percent
- the lower confidence interval limit from the normal approximation results in a value less than 0 percent
- the upper confidence interval limit from the normal approximation results in a value greater than 100 percent.

In any of these circumstances, the Korn and Graubard confidence intervals can and should be asymmetrical.

Confidence intervals for percentiles (eg, medians) are calculated using the Woodruff (1952) method.

Tests for statistically significant differences

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.

Time trends

Where possible, the results of indicators presented in the current report are compared with the corresponding results from the previous years of the continuous NZHS (from 2011/12 onwards) and from the 2006/07 NZHS, to examine whether an indicator shows an increase or a decrease. This is referred to as 'time trends' in the annual report.

Testing the statistical significance of changes over time is based on age-standardised rates.

Section 9: New Zealand Health Survey 2014/15

This section provides some field-related data specific to the data collection and analysis of the NZHS in 2014/15 (year 4). The Appendix contains some information on a past NZHS carried out in 2006/07.

2014/15 NZHS modules

Table 3 outlines the NZHS module topics for 2014/15.

Table 3: NZHS module topics, 2014/15

Adult module topics	Child module topics
Sexual and reproductive health	Child development
Biomedical tests	

The Sexual and Reproductive Health module was a series of questions asked of all adults in the age group of 16–74 years who were cognitively able and did not require a translator.

A subsample of adult respondents was selected and invited to give blood and urine samples, with the aim of achieving 5000 biomedical tests. The subsample selected included all Māori adults, Pacific adults and pregnant women in the NZHS sample.

For further details on the questionnaires for 2014/15, see the *Content Guide 2014/15* (Ministry of Health 2015b). For the survey questionnaire itself, go to www.health.govt.nz .

Data collection

In the second year of the continuous NZHS, 1 July 2014 to 30 June 2015, a total of 13,497 adults and 4754 children took part in the survey. Table 4 shows the number of participants selected in each quarter of 2014/15.
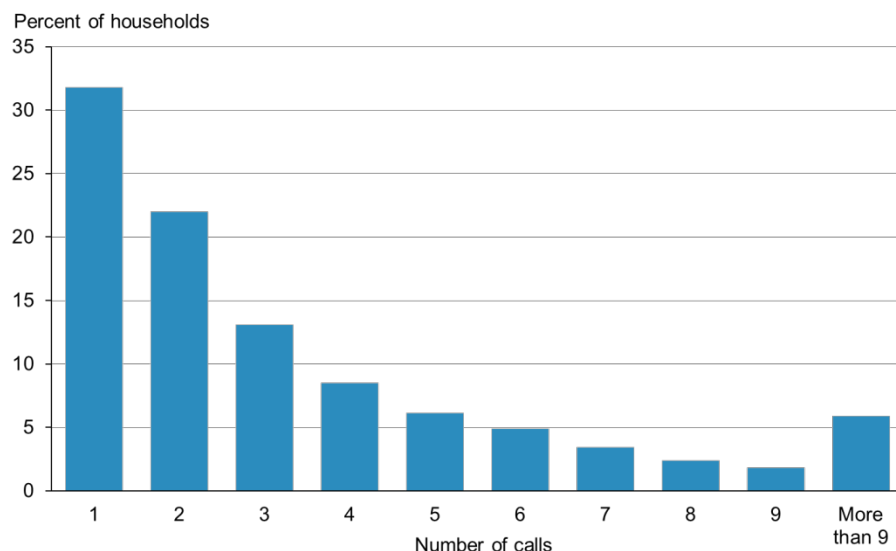
Table 4: Number of survey participants, by quarter, 2014/15

	Adults		Children	
	Number	Percentage of total participants	Number	Percentage of total participants
Quarter 1 (July–September 2014)	3289	24	1123	23
Quarter 2 (October–December 2014)	3433	26	1181	25
Quarter 3 (January–March 2015)	3373	25	1173	25
Quarter 4 (April–June 2015)	3402	25	1277	27
Total (July 2014–June 2015)	13,497	100	4754	100

Call pattern

The call pattern used in the NZHS, as described in Section 4, is an important part of achieving a high response rate. In 2014/15 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).
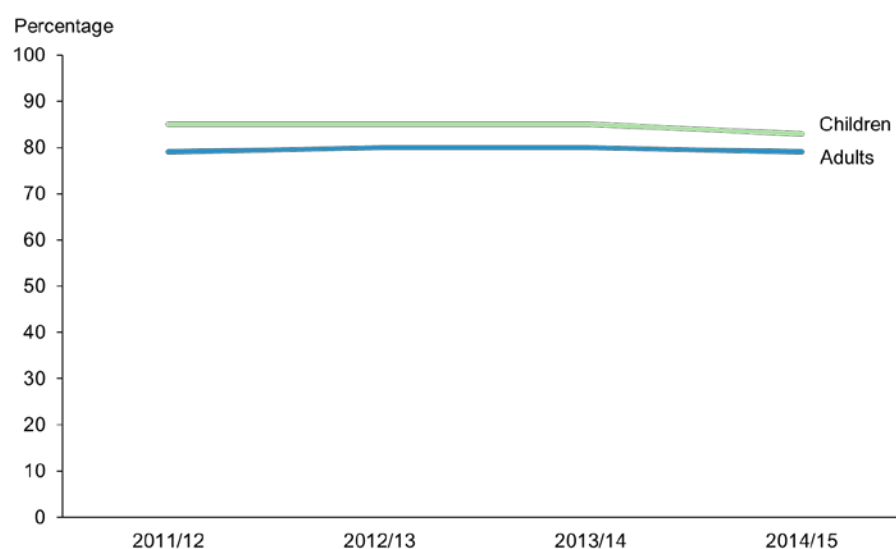
Figure 1: Proportion of households agreeing to first interview, by number of calls, 2014/15

Response rates

The NZHS is well received by the public: the weighted response rates in 2014/15 was 79 percent for adults and 83 percent for children. Figure 2 shows the time trend of response rates of adults and children from 2011/12 to the current survey year, 2014/15.
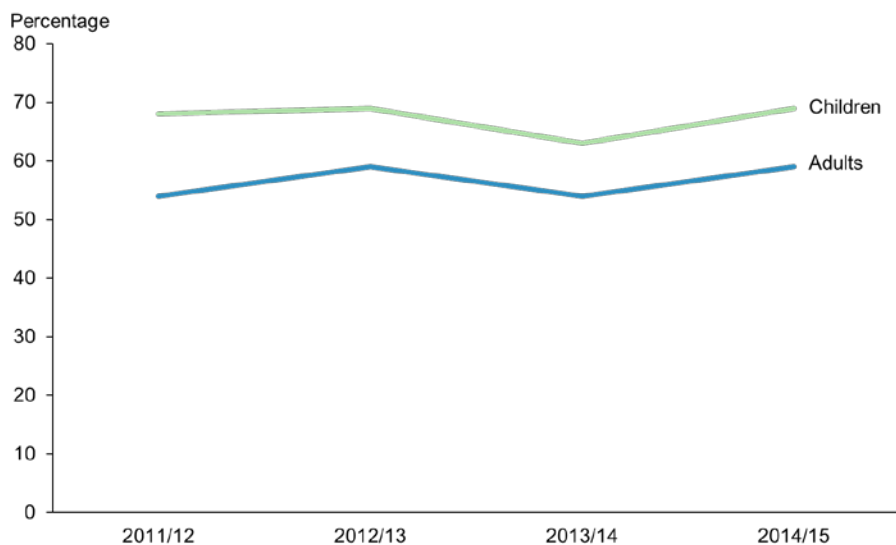
Figure 2: Response rates (%) for adults and children, 2011/12–2014/15

Coverage rates

In 2014/15 the coverage rates were 59 percent for adults and 69 percent for children. Figure 3 shows the time trend of coverage rates of adults and children from 2011/12 to the current survey year, 2014/15.
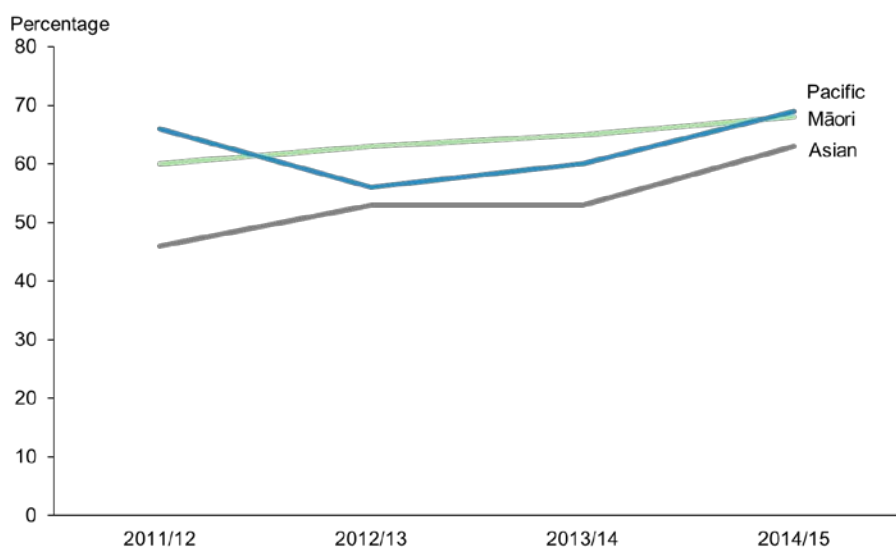
Figure 3: Coverage rates (%) for adults and children, 2011/12–2014/15

Coverage rates were high for children, reflecting high rates for adults in the typical parenting age range.

In 2014/15 the coverage rates were 68 percent for Māori, 69 percent for Pacific peoples and 63 percent for Asian peoples. Figure 4 shows the time trend of coverage rates for Māori, Pacific and Asian ethnic groups from 2011/12 to the current survey year, 2014/15.

Figure 4: Coverage rates (%) for Māori, Pacific and Asian groups, 2011/12–2014/15

As Figure 4 shows, the coverage rates are gradually improving for all three of these ethnic groups – Māori, Pacific and Asian peoples.

In 2014/15 the coverage rates for neighbourhood deprivation quintiles were 53 percent (Q1), 60 percent (Q2), 59 percent (Q3), 60 percent (Q4) and 72 percent (Q5). Figure 5 shows the time trend figures for Q1 to Q5 from 2011/12 to the current survey year, 2014/15.

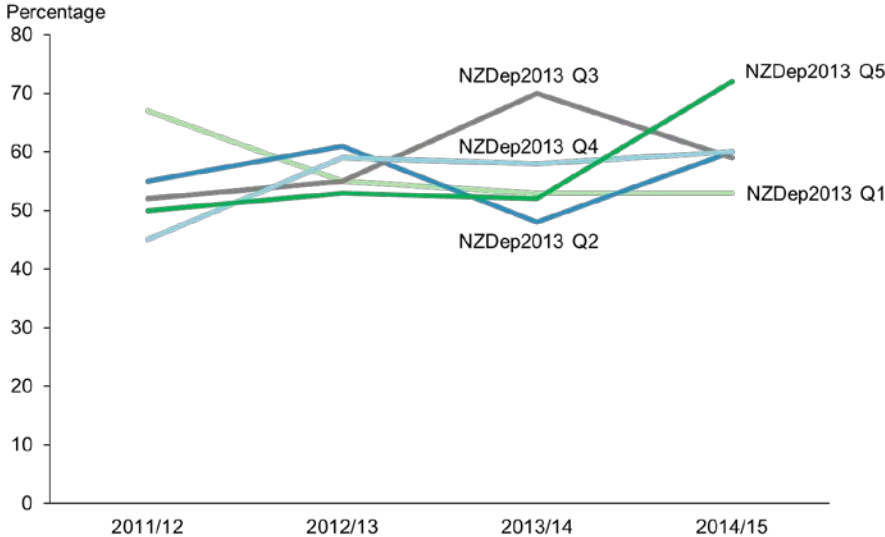**Figure 5: Coverage rates (%) by NZDep2013 quintiles, 2011/12–2014/15**

Figure 6 shows the coverage rates by age and sex for 2014/15. The pattern for Māori in Figure 7 is similar to the overall pattern shown in Figure 6.

Figure 6: Coverage rates (%) for total population, by age group and sex, 2014/15
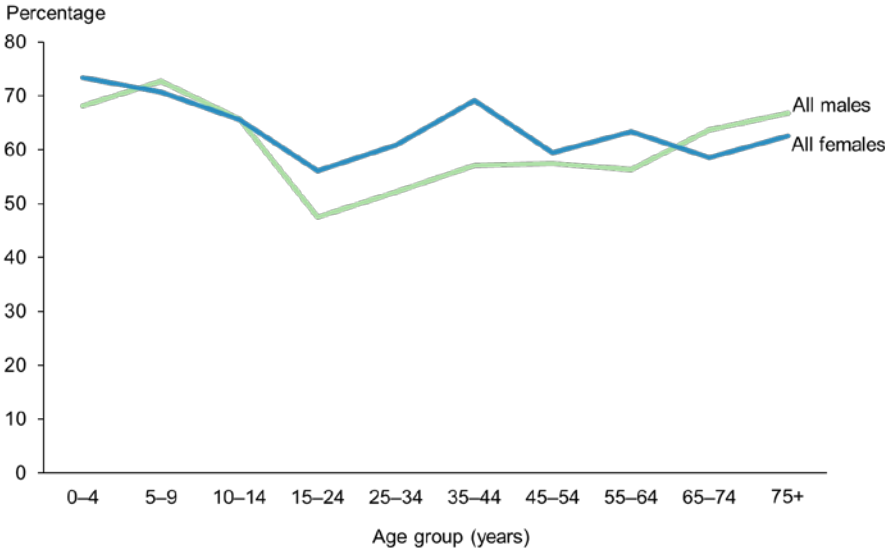
Figure 7: Coverage rates (%) for Māori, by age group and sex, 2014/15

Final weights

Section 7 has explained how the calibrated weights were calculated. Table 5 gives basic descriptive information on the final weights calculated for the 2014/15 survey.

The g-weights are the ratios of the final weights to the initial selection weights. The mean g-weight is approximately 1.9. This means the calibrated weights, which were calculated using population benchmark information, have changed the initial selection weight by an average factor of 1.9.

Table 5: Final weights, 2014/15

	Final weight
Minimum	11
Median	189
90th percentile	576
95th percentile	797
99th percentile	1346
Maximum	1625
Coefficient of variation (CV%)	94.1
Approximate design effect due to weighting $(1 + CV^2)$	1. 9

Sample sizes

Tables 6–9 show the 2014/15 NZHS sample sizes and the total usually resident population counts, by sex, ethnicity, age and NZDep2013 quintile.

Table 6: Sample sizes and population counts for children and adults, by sex, 2014/15

Population group	Sex	Interviews	Measurements* (2+ years)	Population count
Children (0–14 years)	Boys	2421	1882	468,187
	Girls	2333	1833	444,870
	Total	4754	3715	913,057
Adults (15 years and over)	Men	5884	5701	1,772,442
	Women	7613	7068	1,880,210
	Total	13,497	12,769	3,562,652

* Note: These numbers are based on the number of respondents with valid height and weight measurements, and they exclude 201 pregnant women, who are not eligible to be measured.

Table 7: Sample sizes and population counts for children and adults, by ethnic group, 2014/15

Ethnic group (total response)	Population group	Interviews	Measurements* (2+ years)	Population count
European/Other	Children	3061	2412	651,664
	Adults	9729	9242	2,793,262
Māori	Children	1773	1394	234,943
	Adults	3166	2971	474,165
Pacific	Children	689	526	107,198
	Adults	885	820	201,000
Asian	Children	598	427	122,493
	Adults	1174	1109	433,500

* Note: These numbers are based on the number of respondents with valid height and weight measurements, and they exclude 201 pregnant women, who are not eligible to be measured.

Table 8: Sample sizes and population counts, by age group, 2014/15

Age group (years)	Interviews	Measurements* (2+ years)	Population count
0–4	1764	919	306,973
5–9	1484	1401	311,813
10–14	1506	1395	294,273
15–24	1672	1577	650,850
25–34	2125	1952	588,703
35–44	2299	2192	584,068
45–54	2237	2146	631,110
55–64	2159	2074	532,598
65–74	1641	1582	385,702
75 and over	1364	1246	279,622

* Note: These numbers are based on the number of respondents with valid height and weight measurements, and they exclude 201 pregnant women, who are not eligible to be measured.

Table 9: Sample sizes and population counts, by NZDep2013 quintile, 2014/15

NZDep2013 quintile	Population group	Interviews	Measurements* (2+ years)	Population count
Quintile 1 (least deprived neighbourhoods)	Children	517	410	175,365
	Adults	1579	1511	737,777
Quintile 2	Children	729	579	181,594
	Adults	2210	2107	731,548
Quintile 3	Children	762	592	163,222
	Adults	2507	2397	749,920
Quintile 4	Children	1042	814	181,665
	Adults	2964	2808	731,477
Quintile 5 (most deprived neighbourhoods)	Children	1704	1320	211,212
	Adults	4237	3946	701,930

* Note: These numbers are based on the number of respondents with valid height and weight measurements, and they exclude 201 pregnant women, who are not eligible to be measured.

NZDep2013 quintile	Population group	Interviews	Measurements* (2+ years)	Population count

Section 10: Errors in previously published statistics

This section notifies NZHS users about errors in the statistics published in previous annual reports. These errors occurred as a result of independent events at different stages of the survey process, which are explained below.

Revisions to the data and statistics have been made as these errors have been uncovered. The size of the revisions has generally been very small, with the corrected statistics well within the confidence interval limits of the originally published statistics.

Year 1, 2 and 3 estimation weights

The NZHS sample design involves selecting households systematically within meshblocks, using a skip algorithm. In years 1 to 3 of the survey (ie, 2011/12, 2012/13 and 2013/14), the skips used were different from those specified in the design. In particular, households in higher-growth meshblocks had a lower probability of selection than intended.

The estimation weights for respondents in years 1, 2 and 3 had previously been calculated using the intended probabilities of selection. The estimation weight of every respondent in years 1, 2 and 3 has now been recalculated to reflect each respondent's actual probability of selection.

The revisions to the weights also took advantage of the availability of revised population estimates from Statistics New Zealand for the period between 2006 and 2013, which have been re-based to the 2013 Census.

The revisions to the weights mean that all year 1, 2 and 3 indicators in the year 4 report and accompanying tables may be different to those previously published.

An additional issue was that some respondents in the final quarter of year 3 were incorrectly assigned to an NZDep2013 quintile, because they had been surveyed in a 2006 meshblock that was split by Statistics New Zealand into two or more new meshblocks in the 2013 meshblock pattern. These respondents have now been geo-coded to 2013 meshblocks, and the correct NZDep2013 quintile used in both the calibration of the year 3 weights and all year 3 analyses by NZDep2013.

Year 1, 2 and 3 obesity cut-offs

The prevalence statistics for underweight, healthy weight, overweight and obese classes in the 2011/12 and 2012/13 NZHS reports were based on the International Obesity Taskforce (IOTF) cut-off values for body mass index published in Cole et al (2000, 2007).

In 2012, the IOTF slightly revised the cut-off values for those aged 2–18 years (Cole and Lobstein 2012). While the Indicator Interpretation Guide that accompanied the 2013/14 annual report included a table of the revised cut-off values, the 2013/14 prevalence statistics were mistakenly calculated using the old IOTF cut-off values.

The revised IOTF cut-off values have now been used to calculate the prevalence of underweight, healthy weight, overweight and obese classes in all years of the continuous NZHS.

Year 2 height measurements

After publishing the year 2 (2012/13) report, it was discovered that all height measurements for the year 2 respondents had been truncated to the next lowest centimetre. Heights are measured and recorded by CBG's surveyors to the nearest 0.1 cm, but an error had occurred in transferring the data to the Ministry of Health.

Revised height measurements were used in all analyses of year 2 data for the year 3 (2013/14) annual report. Because respondents appeared to be shorter than they really were, the prevalence of obesity reported in year 2 was slightly overstated.

Surveyor dishonesty in year 1

Just before publishing the year 1 (2011/12) report, dishonest practice was discovered in one surveyor's work in the areas of Hutt Valley and Capital & Coast DHBs. The affected data were removed and those households re-interviewed. The year 1 report was published without the data of the re-interviewed households.

Subsequently the re-interviewed records were included in the year 1 data set and the estimation weights recalculated. All year 1 statistics in the annual reports from year 2 onwards are based on the revised year 1 data set.

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.

Bieler GS, Brown GG, Williams RL, et al. 2010. Estimating model-adjusted risks, risk differences, and risk ratios from complex survey data. *American Journal of Epidemiology* 171: 618–23.

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 (accessed 29 October 2015).

Cole TJ, Lobstein T. 2012. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity* 7(4): 284–94.

Cole TJ, Bellizzi MC, Flegal KM, et al. 2000. Establishing a standard definition for child overweight and obesity worldwide: international survey. *British Medical Journal* 320(7244): 1240.

Cole TJ, Flegal KM, Nicholls D, et al. 2007. Body mass index cut offs to define thinness in children and adolescents: international survey. *British Medical Journal* 335(7612): 194.

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

Hayes L, Barry G. 2002. Sampling variability of the Kunst-Mackenbach relative index of inequality. *Journal of Epidemiology and Community Health* 56: 762–5.

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.

Korn EL, Graubard BI. 1999. *Analysis of Health Surveys.* New York: Wiley.

Kott PS. 2001. The delete-a-group jackknife. *Journal of Official Statistics* 17(4): 521–6.

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. 2015a. *Annual Update of Key Results 2014/15: New Zealand Health Survey.* Wellington: Ministry of Health.

Ministry of Health. 2015b. *Content Guide 2014/15: New Zealand Health Survey.* Wellington: Ministry of Health.

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

Statistics New Zealand. 2005. *Understanding and Working with Ethnicity Data: A technical paper.* 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: 2006/07 New Zealand Health Survey

To determine any changes in the prevalence of indicators over time, the annual report, *Annual Update of Key Results 2014/15*, shows results comparing the current NZHS with the previous survey conducted in 2006/07. This appendix gives a brief description of the 2006/07 NZHS.

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 peoples, 5143 Māori, 1831 Pacific peoples and 2255 Asian peoples 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).