

New Zealand Health Survey Methodology Report 2012

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

Published in December 2012 by the
Ministry of Health
PO Box 5013, Wellington 6145, New Zealand

ISBN 978-0-478-40222-3 (online)
HP 5594

This document is available at www.health.govt.nz



Authors

This report was written by Dr Deepa Weerasekera, Dr Niki Stefanogiannis, Robert Templeton and Anne McNicholas (Health and Disability Intelligence Unit, Ministry of Health); Dr Barry Gribben, Carol Boustead and Neil Tee (CBG Health Research Ltd); and Dr Robert Clark (Centre for Statistical and Survey Methodology, University of Wollongong, Australia).

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.

Contents

Section 1: Introduction	1
Background	1
Section 2: The NZHS questionnaire	3
Core component	3
Module component	5
Section 3: Survey population and sample design	6
Target and survey population	6
Sample design	7
Section 4: Data collection	9
Dress rehearsal	9
Enumeration	9
Invitation to participate	9
Call pattern	10
Auditing of interviewers	10
Interviewer training	10
Section 5: Response and coverage rates	11
Calculation of response rate	11
Coverage rates	12
Section 6: Weighting	13
Calculation of selection weights	14
Calibration of selection weights	15
Weights for measurement participants	16
Benchmark populations	16
Calibration software and bounding of weights	18
Section 7: Data processing and analysis	19
Capture and coding	19
Security of information	19
Checking and editing	19
Imputation	20
Creation of derived variables	20
Analysis methods	20

Section 8: New Zealand Health Survey 2011/12	25
2011/12 NZHS module	25
Data collection	25
Response rates	27
Coverage rates	27
Final weights	28
Sample sizes	29

References	32
-------------------	-----------

Appendix 1:	
Comparison of the current NZHS with previous surveys	34
Previous population health surveys	34
Comparability of the surveys	36

List of Tables

Table 1:	Core content of the NZHS	4
Table 2:	NZHS module topics 2011/12 to 2015/16	5
Table 3:	NZHS module topics, 2011/12	25
Table 4:	Number of survey participants by quarter, 2011/12	26
Table 5:	Coverage rates, children and adults combined, 2011/12	27
Table 6:	Final weights, 2011/12	29
Table 7:	Sample sizes and population counts for children and adults, by gender, 2011/12	29
Table 8:	Sample sizes and population counts for children and adults, by ethnic group, 2011/12	30
Table 9:	Sample sizes and population counts, by age group, 2011/12	30
Table 10:	Sample sizes and population counts, by NZDep2006 quintile, 2011/12	31

List of Figures

Figure 1:	Proportion of households agreeing to first interview by number of calls, 2011/12	26
Figure 2:	Coverage rates (%) by age group and gender, 2011/12	28

Section 1: Introduction

The New Zealand Health Survey (NZHS) is an important data collection tool used to monitor population health and provide supporting evidence for health policy and strategy development. The Health and Disability Intelligence Unit 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 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 were 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 will be Tier 1 statistics. This 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). In addition, information specific to the data collection and analysis of the NZHS in 2011/12 are included in section 8.

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 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 cannot be collected more efficiently from other sources, and 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 NZHS content guide contains further information on these objectives (Ministry of Health 2012).

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

- selecting only one eligible adult and one eligible child per dwelling
- using well-tested and largely well-proven questionnaires
- having 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: The NZHS 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 NZHS also includes a set of module topics that change every six or twelve months.

Cognitive testing is undertaken to ensure that 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 NZHS content guide (Ministry of Health 2012).

Core component

The core questions for both adults and children are largely drawn from the main topic areas included in the 2006/07 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 – gender, 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
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	Gender, age, ethnicity, language, country of birth, education, income and income sources, employment status, medical insurance, household characteristics
Anthropometry	Height, weight and waist circumference measurements

Module component

In October 2011 the module topics for the NZHS were agreed for the five years from 2011/12 to 2015/16. These topics 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	To be determined
2015/16	Physical activity	Physical activity

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 3.5 million adults (aged 15 years and over) and 0.9 million children (aged from birth to 14 years), according to the Statistics New Zealand projected population for the year 2011.

Previously the NZHS included only people living in private accommodation. The target population for the current NZHS was extended to particularly include 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, those located off the main islands of New Zealand (North, South and Waiheke) and others on other sparsely inhabited off-shore islands, on-shore islands, waterways and inlets.

Sample design

The sample design for the NZHS has been developed by the Centre for Statistical and Survey Methodology, 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).

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 where 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 the 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. The sample has been initially selected for a period of 12 quarters (three years, 2011 to 2013). Second, a list of households is compiled for each selected meshblock. An equal probability sample of 20 households is selected from this list. Finally, one adult (aged 15 years or over) and one child (aged from birth to 14 years old, 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 old, 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 as 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 typing responses directly into a laptop computer using 'Survey System' computer assisted personal interview (CAPI) software. Showcards with predetermined response categories are used to assist respondents, where appropriate.

Dress rehearsal

A dress rehearsal was carried out from 31 March to 17 April 2011 in 20 meshblocks in Northland, Auckland and Waikato. 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 and the full survey was extended across all of New Zealand during May and June 2011. There was a small delay in starting recruitment activities in Christchurch because of the impact of the 22 February 2011 earthquake.

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 this 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 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 gender 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 Health Survey or 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. Where contact has not been made already, six calls are made in the survey month in which the meshblock is issued. 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 5 percent of all participants or 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 on-going training courses run by CBG on how to conduct interviews.

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 2011/12 the final weighted response rates were 79 percent for adults and 85 percent for children.

For more details on the response rate for 2011/12, 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:

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

Coverage rates

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 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 (for example, 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 the response by 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 2011/12, see section 8.

Section 6: Weighting

Weighting of survey data ensures that the estimates calculated from these data are representative of the target population.

Most national surveys have complex sample designs where 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 that the sample is geographically clustered, or 'clumped'
- ensure that all regions of interest, including small regions, have sufficient sample to enable adequate estimates
- ensure that all sub-populations, in particular the Māori, Pacific and Asian populations, have sufficient sample to enable adequate estimates.

To ensure that 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 who 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 would receive lower weights. Also, groups who have a lower response rate (for example, 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 to 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.

Although, the selection weights are used only as inputs to the calibration process, it is crucial to calculate them correctly because:

- the bias of estimates based on calibrated weights depends entirely on the correct selection weights; if selection weights do not represent the probability of selection, then calibrated estimators will be biased
- the calibration process allows the selection weights to correspond to population benchmarks.

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 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 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, and $E_{ij} = 0$ 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 current 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:

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

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

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

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

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

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

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 the population counts from the 2006 Census broken down by age, sex, district health board (DHB) and ethnicity 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 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, p 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:

$$\text{final weight} = \text{initial weight} * \text{g-weight}.$$

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
- 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 (NZDep2006) quintile.

Age, sex, ethnicity and socioeconomic position (Māori, Pacific, Asian self-identified total ethnicity and NZDep2006) 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 NZDep2006 quintiles

Benchmarks for the quintiles of the New Zealand Index of Deprivation 2006 (NZDep2006) are derived by dividing the total population estimates (all ages) into five groups of equal size. This approach follows the original construction of the NZDep quintiles, which divide the total population into five equal-sized groups (children and adults combined). Although the NZDep index is based on 2006 Census data and the proportions of the population within each quintile will have changed since then, an analysis of the quarterly Household Labour Force Survey data (covering the period 2006 to 2012) showed that the level of change is not significant. For this reason, it was decided to use these NZDep2006 quintile benchmarks to help control quarterly variation in the NZHS sample distribution.

Comparison with benchmarks used in the 2006/07 NZHS

The 2006/07 NZHS used more extensive benchmarks, but this is not feasible in the current NZHS. The following are some particular differences between the two surveys.

- In 2006/07 Māori benchmarks were broken down by the same age grouping as total benchmarks. In the current NZHS, these categories are broadened, with wider 15–29 and 55–64 age groups and with top-coding at 65+ years. This is because weights will be calculated for each quarterly sample of the current NZHS, so that the total sample size is only around 3000 adults per quarter, rather than the approximately 12,000 adults in the 2006/07 survey. Because of this focus on the quarterly sample, the finer age grouping resulted in too many small cells, making weights unstable.
- District health board benchmarks (by child/adult) were used in 2006/07 but will not be used in the current NZHS. Statistics New Zealand releases annual population estimates by DHB, although not broken down by child and adult. These population estimates are compiled by updating the most recent Census (2006) for age, sex, international migration and internal migration. At detailed regional levels, such as DHB areas, these estimates may be less reliable when it has been some time since the most recent Census, as is presently the case. DHB benchmarks were more feasible in 2006/07 because the Census data were only one year out of date so that Census counts updated by simple pro-rating were thought to be adequate.
- Benchmarks by Pacific and Asian ethnicity were used in 2006/07 based on Census data. At that time the Census data were only one year out of date so that it was feasible to use pro-rated Census counts. This would not be reliable in the current NZHS as the most recent Census data are five to eight years out of date. For the quarters of the ongoing NZHS survey, benchmarks derived from the Household Labour Force Survey (as discussed above) are used.

- The New Zealand Index of Deprivation was considered as a possible benchmark variable in 2006/07. The weighted counts from the NZHS sample by NZDep were calculated, and the distribution reflected the 2006 Census distribution quite closely without explicitly using NZDep in weighting. At that time, therefore, it was decided that NZDep should not be used in weighting. For the current NZHS, there is more variability in the sample distribution across NZDep quintiles. Because weighting is done in each quarter, NZDep benchmarking is worthwhile. An analysis of the Household Labour Force Survey data showed that the distribution of the population across the quintiles has been stable since 2006. Therefore making a simple allocation of the total population equally across the quintiles is a valid way to create the required benchmarks.

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 6500 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. Trial and error showed that bounding the g-weight resulted in either no weights being truncated or a very large number. This finding reflects the requirement that g-weights are a function of age group, sex and Māori identification only and so there is a relatively small number of classes, with equal g-weights within class. Bounding the g-weight means that either no records or whole classes are affected.

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 could 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 security login username and password.

Names and addresses of people and households who 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.

Imputation

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 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.

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 district health board).

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 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 in 2013.

Analysis methods

The data are analysed according to the following techniques for NZHS reports.

Calculation of proportions

The proportion of the population who belong to a particular group (for example, the proportion of the population who have diabetes) is estimated by: 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 (for example, the proportion of Māori who have diabetes) is estimated by: 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).

Calculation of totals (counts)

Estimates of totals are given by the sum over 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 number of respondents with diabetes) multiplied by the weight.

Calculation of averages (means)

Estimates of the population averages (for example, the average number of visits to a general practitioner or GP) are calculated by: 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: 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.

Ethnicity

The reports use **total response ethnicity** to define ethnic groups. Total response ethnicity classifies a person in all the ethnic groups that 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 2006 (NZDep2006), which measures the level of socioeconomic deprivation for each neighbourhood (meshblock) according to a combination of the following 2006 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 (Salmond et al 2007).

Survey data are generally presented for NZDep2006 quintiles 1–5. Quintile 1 represents the 20 percent of small areas with the lowest levels of deprivation (least deprived areas) and quintile 5 represents the 20 percent of small areas with the highest level of deprivation (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 most and least deprived points on the regression line are used to calculate the relative index of inequality.

Age standardisation

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 gender, ethnic group and neighbourhood deprivation (NZDep2006), 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.

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.

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 general, comparisons refer to the 'group of interest' (eg, men) compared with the 'reference group' (eg, women).

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

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 the true value of the estimate (if we were to survey the whole population) lies between the lower and upper confidence interval values.

Differences between estimates are said to be statistically significant when the confidence intervals for each rate do not overlap. Sometimes, however, even when there are overlapping confidence intervals, the difference between the groups can be statistically significant. Any differences between two variables where the confidence intervals overlap are tested using a *t*-test. The significance of a *t*-test is represented by the *p*-value. If a *p*-value is below 0.05, then we are 95 percent confident the difference between the two estimates is statistically significant.

Percentiles

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

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 that 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 symmetric 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 be asymmetrical.

Adjusting population totals for item non-response

To account for item non-response in population total estimates, a factor is calculated using the sum of the weighted denominator and the weighted number of item non-respondents divided by the weighted denominator. This is applied to both the weighted numerator and the weighted denominator.

Section 8: New Zealand Health Survey 2011/12

This section provides some field-related data specific to the data collection and analysis of the NZHS in 2011/12. Appendix 1 contains information on the comparability of the current NZHS with previous surveys.

2011/12 NZHS module

Table 3 outlines the NZHS module topics for 2011/12.

Table 3: NZHS module topics, 2011/12

Adult module topics	Child module topics
Health service utilisation	Health service utilisation
Patient experience	Patient experience
Problem gambling	
Discrimination	

Further details on the questionnaires for 2011/12 can be found in the *New Zealand Health Survey Content Guide 2011–2012* (Ministry of Health 2012). The survey questionnaire can be found at www.health.govt.nz.

Data collection

In the first year of the continuous NZHS, 1 July 2011 to 30 June 2012, a total of 12,370 adults and 4478 children took part in the survey.

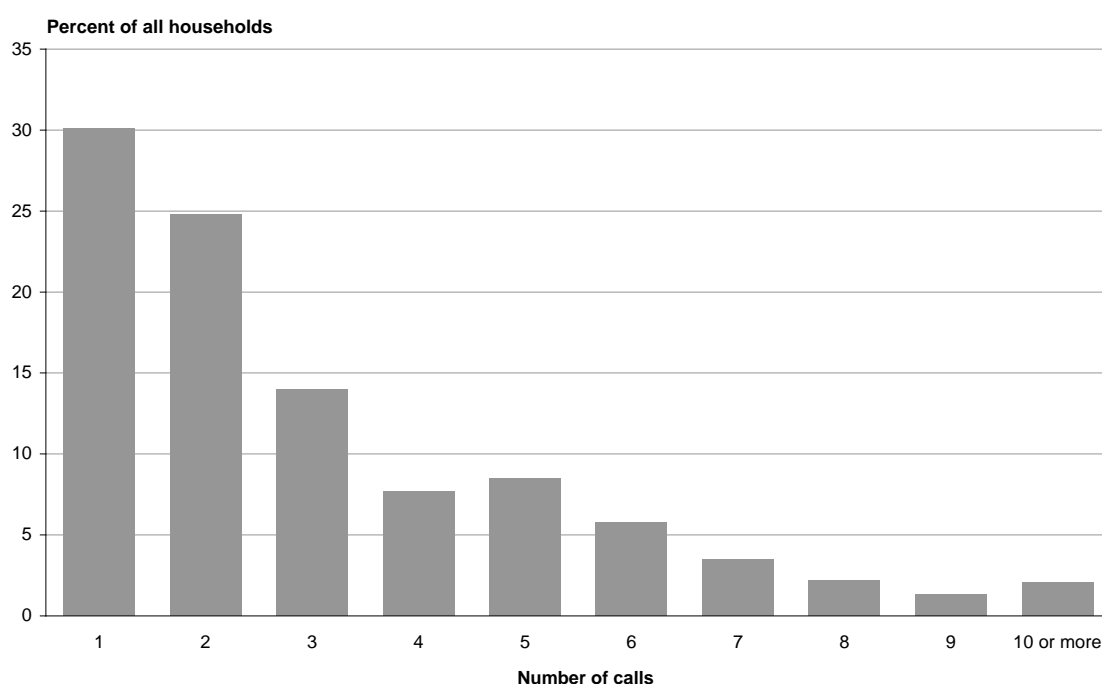
Table 4 shows the number of participants selected each quarter in 2011/12.

Table 4: Number of survey participants by quarter, 2011/12

	Adults		Children	
	Number	Percent of total participants	Number	Percent of total participants
Quarter 1 (July–September 2011)	3041	25%	1129	25%
Quarter 2 (October–December 2011)	3209	26%	1082	24%
Quarter 3 (January–March 2012)	3251	26%	1237	28%
Quarter 4 (April–June 2012)	2869	23%	1030	23%
July 2011–June 2012	12,370	100%	4478	100%

Call pattern

The call pattern used in the NZHS was an important component of achieving a high response rate. Surveyors followed a proven call approach including visiting meshblocks at different times and on different days depending on the area that they were working in. For 94 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, 2011/12

Response rates

The weighted response rates in 2011/12 were 79 percent for adults and 85 percent for children.

These response rates are much higher than the response rates of the previous 2006/07 NZHS (68 percent for adults and 71 percent for children). They indicate that the current NZHS was well received by the public.

The collection of measurement data (including height and weight) had a slightly lower response rate among survey respondents than in previous surveys. The measurement data were reweighted so that they represented the total population. Because the current survey had a higher overall response rate for the whole survey, the overall response rate for the measurement collection is comparable across survey years.

Coverage rates

In 2011/12, the coverage rate was 54 percent for adults and 68 percent for children.

Table 5 shows the coverage rates in 2011/12 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 declined with higher levels of neighbourhood deprivation, which is typical of national surveys. The effect was removed through the use of calibrated weights.

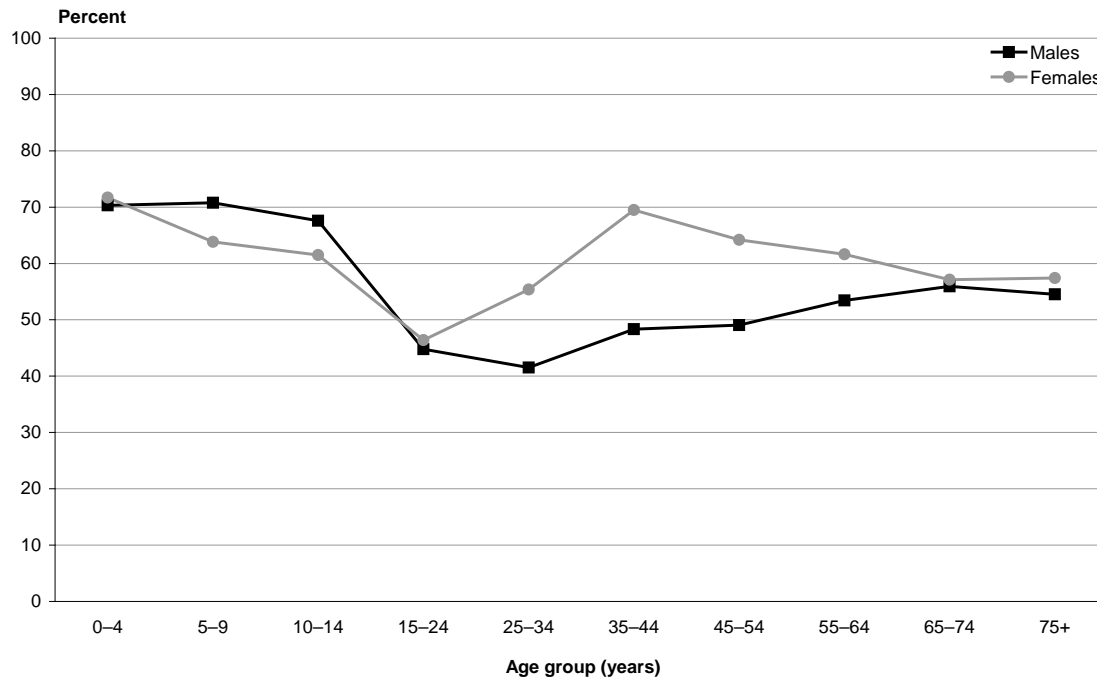
Coverage rates were high for children, reflecting high rates for adults in the typical parenting age range (Figure 2).

Table 5: Coverage rates, children and adults combined, 2011/12

Population	Coverage rate (%)
Māori	60
Pacific	66
Asian	46
NZDep2006 Quintile 1	67
NZDep2006 Quintile 2	55
NZDep2006 Quintile 3	52
NZDep2006 Quintile 4	45
NZDep2006 Quintile 5	50

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

Figure 2: Coverage rates (%) by age group and gender, 2011/12



Final weights

Section 6 explained how the calibrated weights were calculated. Table 6 gives basic descriptive information on the final weights calculated for the 2011/12 survey.

The ratio (the g-weights) of the final weight to the initial selection weight is 1.92. That is, the calibrated weights, which were calculated using population benchmark information, have changed the initial selection weight by a factor of 1.92.

Table 6: Final weights, 2011/12

	Final weights
Minimum	10
Median	210
90th percentile	591
95th percentile	774
99th percentile	1271
Maximum	1625
Coefficient of variation (CV%)	86.20
Approximate design effect due to this component of weighting (1+CV ²)	1.74

Sample sizes

Tables 7 to 10 show the 2011/12 NZHS sample sizes and the total usually resident population counts by gender, ethnicity, age and NZDep2006 quintile.

Table 7: Sample sizes and population counts for children and adults, by gender, 2011/12

Population group	Sex	Interviews	Measurements (2+ years)	Population count
Children (0–14 years)	Boys	2309	1588	457,000
	Girls	2169	1459	435,408
	Total	4478	3047	892,408
Adults (15 years and over)	Men	5014	4509	1,717,923
	Women	7356	6015	1,814,133
	Total	12,370	10,524	3,532,055

Table 8: Sample sizes and population counts for children and adults, by ethnic group, 2011/12

Ethnic group (total response)	Sex	Interviews	Measurements (2+ years)	Population count
European/Other	Children	2968	2056	630,947
	Adults	9315	8028	2,783,286
Māori	Children	1592	1009	230,336
	Adults	2543	2072	444,224
Pacific	Children	730	492	142,654
	Adults	938	780	205,176
Asian	Children	432	292	88,126
	Adults	915	793	374,647

Table 9: Sample sizes and population counts, by age group, 2011/12

Age group	Interviews	Measurements (2+ years)	Population count
0–4 years	1790	761	313,383
5–9 years	1267	1091	289,303
10–14 years	1421	1195	290,423
15–24 years	1464	1264	643,175
25–34 years	1923	1597	572,678
35–44 years	2330	2024	596,798
45–54 years	2088	1830	618,345
55–64 years	1928	1662	498,888
65–74 years	1384	1181	336,955
75 years and over	1253	966	265,218

Table 10: Sample sizes and population counts, by NZDep2006 quintile, 2011/12

NZDep2006 quintile	Sex	Interviews	Measurements (2+ years)	Population count
Quintile 1 (least deprived neighbourhoods)	Children	643	486	155,084
	Adults	1932	1698	729,949
Quintile 2	Children	671	466	163,623
	Adults	1948	1677	721,410
Quintile 3	Children	829	549	173,220
	Adults	2459	2103	711,812
Quintile 4	Children	915	605	182,442
	Adults	2652	2292	702,591
Quintile 5 (most deprived neighbourhoods)	Children	1420	941	218,739
	Adults	3379	2754	666,293

References

- Ahmad O, Boschi-Pinto C, Lopez A, et al. 2000. *Age-standardization of rates: A new WHO standard*. Geneva: World Health Organization.
- Blakely T, Tobias M, Atkinson J, et al. 2007. *Tracking Disparity: Trends in ethnic and socioeconomic inequalities in mortality, 1981–2004*. Wellington: Ministry of Health.
- 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.
- 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. 1999. *Taking the Pulse: The 1996/97 New Zealand Health Survey*. Wellington: Ministry of Health.
- Ministry of Health. 2003. *NZ Food, NZ Children: Key results of the 2002 National Children's Nutrition Survey*. Wellington: Ministry of Health.
- Ministry of Health. 2004. *A Portrait of Health: Key results of the 2002/03 New Zealand Health Survey*. Wellington: Ministry of Health.
- 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. *The New Zealand Health Survey Content Guide 2011–2012*. Wellington: Ministry of Health.
- Robson B, Harris R. 2007. *Hauora: Māori Standards of Health IV: A study of the years 2000–2005*. Wellington: Te Rōpū Rangahau Hauora a Eru Pōmare, School of Medicine and Health Sciences, University of Otago.
- Russell DG, Parnell WR, Wilson NC, et al. 1999. *NZ Food: NZ People: Key results of the 1997 National Nutrition Survey*. Dunedin: LINZ® Activity and Health Research Unit, University of Otago.
- Salmond C, Crampton P, Atkinson J. 2007. *NZDep2006 Index of Deprivation User's Manual*. Wellington: Department of Public Health, University of Otago.

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

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: Comparison of the current NZHS with previous surveys

To determine any changes in the prevalence of indicators over time, some analyses are carried out comparing the current NZHS with previous surveys.

As a population group, Māori have the poorest health outcomes and highest mortality rate of any ethnic group in New Zealand (Blakely et al 2007; Robson and Harris 2007). It is therefore important to monitor progress towards addressing these disparities. For this reason, time trend data for the Māori population by gender will be included in any time trend analyses where possible.

European/Other, Pacific and Asian people can also have poor outcomes for some indicators. Time trends for these ethnic groups are also explored in survey publications.

Where possible, data from the current NZHS have been compared with data from the 1996/97, 2002/03 and 2006/07 NZHS as well as the 1997 National Nutrition Survey and the 2002 National Children's Nutrition Survey where appropriate.

This section gives a brief description of the surveys used in time trend analyses and provides information on the comparability of the surveys. Survey comparisons with the 1992/93 NZHS have not been carried out due to the different mode of data collection used in that survey (telephone interviews).

Previous population health surveys

1996/97 New Zealand Health Survey

The target population for the 1996/97 NZHS was defined as the total usually resident, civilian population of New Zealand of all ages, residing in permanent private households.

Statistics New Zealand followed a stratified cluster sampling process to select a sample from the target population. The sampling frame was area-based using Statistics New Zealand primary sampling units. Māori and Pacific peoples were oversampled in order to obtain more reliable estimates. There was also some regional oversampling.

Data were collected from October 1996 to October 1997 using face-to-face interviewing. The adult response rate was 74 percent. The final sample was made up of 7862 adults (including 1321 Māori adults) and 1019 children (including 231 Māori children).

For full details on the methodology of the 1996/97 NZHS, see *Taking the Pulse: The 1996/97 New Zealand Health Survey* (Ministry of Health 1999).

1997 National Nutrition Survey

The 1997 National Nutrition Survey (1997 NNS) used the 1996/97 NZHS sample population as a frame from which to recruit participants. Data for the 1997 NNS were collected over the 12-month period from December 1996 to November 1997. The 1997 NNS achieved a response rate of 50 percent, taking into account the response rate of the 1996/97 NZHS. A total of 4636 adults participated in the 1997 NNS, including 704 Māori adults.

For full details on the methodology of the 1997 NNS, see *NZ Food: NZ People: Key results of the 1997 National Nutrition Survey* (Russell et al 1999).

2002 National Children's Nutrition Survey

The target population for the 2002 National Children's Nutrition Survey (CNS) was children aged 5–14 years from schools throughout New Zealand. A sample of 160 schools was selected from the Ministry of Education list. Schools on the Chatham Islands, correspondence schools, and schools with fewer than 50 students were excluded. Te Kura Kaupapa schools were included regardless of size. Students were then randomly selected from the rolls of the selected schools in proportion to the number of students on the school roll. Different sampling proportions from the school roll were used to ensure approximately equal numbers of Māori, Pacific and European/Other children in the survey sample.

Data were collected from February to December 2002. The anthropometric measurements were carried out at school, and the main interview was carried out in the home in the presence of a parent or caregiver. The overall response rate for schools was 91 percent. The response rate of children from participating schools was 69.3 percent. A total of 3275 children participated in the survey, including 1224 Māori children.

For full details on the methodology of the CNS, see *NZ Food, NZ Children: Key results of the 2002 National Children's Nutrition Survey* (Ministry of Health 2003).

2002/03 New Zealand Health Survey

The target population for the 2002/03 NZHS was the usually resident New Zealand adult population, aged 15 years and over, living in permanent private dwellings. An area-based frame using meshblocks as primary sampling units was used as the sample frame. Māori, Pacific and Asian peoples were oversampled.

Data were collected from September 2002 to January 2004 using face-to-face interviewing. The total response rate for the survey was 72 percent. A total of 12,929 people responded to the survey, including 4369 Māori.

For full details on the methodology of the 2002/03 NZHS, see *A Portrait of Health: Key results of the 2002/03 New Zealand Health Survey* (Ministry of Health 2004).

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 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).

Comparability of the surveys

Target population

The 1996/97, 2002/03 and 2006/07 New Zealand Health Surveys had the same target population. The 2011/12 NZHS target population was extended to include those living in aged-care facilities and those temporarily living away from the household in student accommodation.

The target population for the 2002 CNS was children aged 5–14 years from schools throughout New Zealand, excluding Chatham Islands, correspondence schools, and schools with fewer than 50 students (except for Te Kura Kaupapa).

Survey design

Māori, Pacific and Asian peoples were oversampled in the 2002/03 and 2006/07 surveys in order to provide more reliable results. Because there was no oversampling in 1996/97, this survey had a smaller sample size for Māori.

Instead of oversampling, the current NZHS uses two strategies to boost the sample of specific ethnic groups.

1. The electoral roll sample is used to boost the Māori sample size.
2. The area-based sample targets Māori, Pacific and Asian ethnic groups by assigning higher probabilities of selection to meshblocks with higher concentrations of these groups.

The above two strategies have replaced the 2006/07 NZHS approach of proxy household screening for ethnicity.

The smaller sample size for Māori in the 1996/97 NZHS has to be taken into account when comparing variables for this ethnic group in the time trend analyses.

Data collection

The majority of interviews for the 1996/97, 2006/07 and 2011/12 Health Surveys and the 1997 NNS were completed over a period of one year.

The 2002 CNS collected data over 11 months (February to December) and the 2002/03 NZHS collected data over 16 months.

Statistics New Zealand collected the data for the 1996/97 NZHS. The University of Otago carried out the 1997 NNS, and it also collected data for the 2002 CNS, along with the University of Auckland and Massey University (Palmerston North). Data collection for the 2002/03 and 2006/07 Health Surveys was undertaken by National Research Bureau Ltd (NRB). CBG Health Research Ltd (CBG) collected the data for the 2011/12 NZHS.

Different practices between the various organisations may have introduced some unknown bias into the results, in particular when comparing the 1996/97 data with later data. For example, to control for non-response, Statistics New Zealand allowed for four to six call-backs before accepting a dwelling as a non-contact dwelling, whereas NRB and CBG allowed up to 10 call-backs.

Response rates

The 1996/97, 2002/03 and 2006/07 Health Surveys had similar response rates. In 2011/12 the NZHS achieved higher response rates (79 percent for adults and 85 percent for children) than the previous surveys. However, there is no evidence to believe that this difference has any impact on the comparability of the overall results.

The 1997 NNS achieved a response rate of 50 percent, taking into account the response rate of the 1996/97 NZHS.

The 2002 CNS achieved an overall response rate of 91 percent for schools. The response rate of children from participating schools was 69.3 percent.

Analysis

Calibrated weighting

To make the current and past surveys more comparable, the weights from the earlier surveys have been re-benchmarked, using benchmarks that reflect the target population of the current survey.

Comparability of questions

Analyses exploring time trends will be restricted to questions that are comparable between the surveys. Although care will be taken to ensure that only questions with similar wording are used to assess changes in indicators, caution is required when comparing the results as other factors (such as question order) can influence responses to an unknown extent.

Body size calculation

In calculating body mass index, the 1997 NNS and 2002 CNS height and weight data were re-calculated using the same method as used in the 2002/03 and 2006/07 data sets. For the 2002 CNS, and 2006/07 and 2011/12 health surveys, the International Taskforce on Obesity BMI cut-offs were used to define thinness, overweight and obesity using exact age for participants aged 2–17 years (Cole et al 2000, 2007). For the 1997 NNS and 2002/03 NZHS, age was only available in years, so this was used in calculating the body size classifications for those aged 15–17 years. It is unlikely that using years rather than exact age in calculating body size for participants in this age group will have any impact on the overall time trends for adults, by gender or ethnic group.