# Chapter 1: Introduction and Methodology

## Introduction

This chapter overviews the 2002/03 New Zealand Health Survey (NZHS), including its aims, background, content and methodology, and the dissemination of its results.

### Aims

The aims of the 2002/03 NZHS were to:

- measure the health status of New Zealand adults, including their self-reported physical and mental health status, and the prevalence of selected health conditions
- measure the prevalence of risk and protective factors associated with these health conditions
- measure the use of health services, including satisfaction with health services and barriers to accessing health services
- examine differences between population subgroups (as defined by sex, ethnicity, age and the New Zealand Deprivation Index 2001 (NZDep2001))
- examine changes over time.

### Background

The 2002/03 NZHS is part of the *New Zealand Health Monitor* (Ministry of Health 2002), a programme of population health surveys used by the Ministry of Health to monitor New Zealanders' health.

The 2002/03 NZHS was the third national health survey of New Zealanders aged 15 years and over. The two previous surveys were the 1992/93 Household Health Survey (Statistics New Zealand and Ministry of Health 1993) and the 1996/97 New Zealand Health Survey (Ministry of Health 1999). The 1992/93 survey involved telephone interviews with 7065 adults. The 1996/97 survey involved face-to-face interviews with 7862 adults. The 2002/03 NZHS involved face-to-face interviews with 12,929 adults.

All people aged 15 years and over who were usually resident within permanent private dwellings were eligible for selection in the 2002/03 NZHS. In addition, a separate study of people living in institutions was undertaken.

A key objective of the 2002/03 NZHS was to improve the quality of ethnic estimates. Māori, Pacific and Asian ethnic groups were over-sampled to provide more reliable results. As a result, more Māori, Pacific and Asian people participated than ever before. The Chatham and Pitt Islands were also over-sampled.

The 2002/03 NZHS was similar to the two previous national health surveys, particularly the 1996/97 survey. However, the 2002/03 NZHS differs from the previous survey by including a wider range of chronic conditions and risk and protective factors associated with these conditions, additional measures of self-reported physical and mental health status, a wider range of health care services, and height, weight and waist circumference measurements.

Statistics New Zealand developed and tested the 2002/03 NZHS questionnaire. National Research Bureau Ltd (NRB) was contracted to field the 2002/03 NZHS questionnaire. NRB's role included designing the sampling methodology, piloting the questionnaire, obtaining ethical approval, interviewing respondents, processing data and providing a dataset with appropriate documentation to the Ministry of Health.

## Survey content

The 2002/03 NZHS has four health-related modules and a sociodemographic module (Table 1).

**Table 1:** Content of 2002/03 New Zealand Health Survey

| Module | Topics | Details |
|---|---|---|
| Chronic disease | Heart disease, stroke, diabetes, asthma, chronic obstructive pulmonary disease, arthritis, spinal disorders, osteoporosis, cancer, other long-term illnesses. | Prevalence, age at diagnosis, treatments. |
| Health service use | Māori health providers, Pacific health providers, general practitioners, medical specialists, nurses, pharmacists and prescriptions, complementary and alternative medicine providers, other health providers, telephone and internet helplines, hospitals. | Frequency of contact, reasons for visit, satisfaction levels and reasons for dissatisfaction, unmet need and barriers to access. |
| Risk and protective factors | High blood pressure, high blood cholesterol, overweight and obesity, physical activity, tobacco smoking, marijuana smoking, vegetable and fruit intake, alcohol use, gambling. | Prevalence. |
| Self-reported health status | General heath, vision, hearing, digestion, breathing, pain, mental health, sleep, energy and vitality, understanding and remembering, communicating, physical functioning, self-care, usual activities, social functioning. | SF-36 Health Status Questionnaire embedded within the World Health Organization Long Form Health Status Questionnaire. |
| Sociodemographic | Age, sex, ethnicity and responses to ethnicity, country of birth, household characteristics, education, income support, employment, income, medical insurance, NZDep2001 (from meshblock). | |

# Methodology

The following discussion explains the survey design methodology for the main survey, which includes the Chatham and Pitt Islands over-sample, and the separate institutions survey. These surveys were conducted from September 2002 to January 2004.

## Population and frame

### Target population

For the main survey the target population was the New Zealand adult population (ie, people aged 15 years and over) living in permanent private dwellings. The target population was approximately 2.6 million people according to the March 2001 New Zealand Census of Population and Dwellings (2001 Census).

For the institutions survey the target population was all people aged 15 years and over who were resident with relative permanency in hospitals and IHC and rest homes, or who were 'dependent persons'. The 'dependent persons' group refers to people primarily released into communities during the phasing out of the large mental health residential institutions. This group may include people with mental health, drug or alcohol problems, or the residual effects of these.

### Survey population

#### Main survey

**Geographic coverage**

For practical reasons a few households that were part of the defined target population were excluded from participating in the survey, but were accounted for in the final estimates via the survey weights. These included households not resident on the main islands of New Zealand (North, South and Waiheke) and the Chatham and Pitt Islands, such as those located on other off-shore islands, on-shore islands, waterways and inlets.

**Dwellings coverage**

The survey covered the eligible population living within permanent, private dwellings. Private dwelling types that were not included in the survey were temporary private dwellings such as caravans, cabins or tents in a motor camp, or boats. All non-private dwellings were excluded from the survey such as hotels, motels, guest houses, boarding houses, homes for older people, hostels, motor camps, hospitals, barracks and prisons.

**Eligible respondents**

All people aged 15 years and older who were usually resident within permanent private dwellings were eligible for selection as respondents. The term 'usually resident' excluded people who were present within the dwelling at the time of interview but who usually resided elsewhere (either within New Zealand or overseas).

### Institutions survey

The survey population consisted of all persons aged 15 years and over who were resident at the time of the survey with relative permanency in health-related residential establishments known by the Ministry of Health.

## Sample frame

### Main survey

The survey frame was an area-based frame. The frame was the list of small geographic areas (meshblocks) defined by Statistics New Zealand that fall within the geographical coverage of the survey. Meshblocks were the primary sampling units (PSUs).

The survey frame provides the first stage in the sampling process that proceeds to dwelling selection within the meshblock and then respondent selection within the dwelling. The procedure for this selection process is described in the sample design section.

All New Zealand households were geographically clustered to avoid having to list and maintain a frame of all households in the country. Lists of dwellings were only enumerated within selected PSUs. The cost of interviewing was reduced since the selected households were geographically clustered and travelling costs were reduced within clusters.

Meshblocks vary in both population and area size. When the New Zealand-wide standard meshblock system was established in 1976, urban meshblocks contained an average of 150–200 people, while rural meshblocks generally had 100–150 people. Subsequent division of meshblocks and changes in population patterns have resulted in meshblock population counts having considerable range. Some meshblocks have a nil population and some contain more than 500 people.

Ideally, the PSUs would be defined to be less variable in terms of the number of in-scope persons. However, to develop such a frame for this iteration of the survey was not viable due to cost and timeliness constraints.

### Institutions survey

The overall frame was made up of four independent non-overlapping frames provided by the Ministry of Health. For each frame there was a list of establishments, along with contact details and the number of residential beds. The four frames were hospitals (330 establishments), IHC homes (345), rest homes (506) and dependent persons (524).

The hospital frame, which initially contained 444 elements, was used in a preliminary census to establish the number of long-term residential beds in use. The results showed that 330 of the 444 establishments in the initial frame had at least one such bed. The remaining three frames were accepted as accurate with respect to the number of occupied residential beds.

The final frame consisted of 1705 establishments containing 36,698 beds (Table 2).

**Table 2:** Institutions survey frame

| Frames | Number of establishments | Number of residential beds |
|---|---|---|
| Hospitals | 330 | 15,138 |
| IHC homes | 345 | 1,859 |
| Rest homes | 506 | 15,494 |
| Dependent persons | 524 | 4,207 |
| Total | 1705 | 36,698 |

## Sample design

### Overview

#### Main survey

The main survey used a complex sample design to provide high quality estimates for minimal cost and acceptable respondent burden. Population characteristics from the 2001 Census were used in the sample design and sample selection.

All New Zealand dwellings were geographically clustered to avoid having to list and maintain a frame of all dwellings in the country. Lists of dwellings were then enumerated and sampled within selected geographic areas. The cost of interviewing was reduced since the selected households were geographically clustered and travelling costs were reduced within clusters. To minimise respondent burden only one eligible person was selected from each sampled dwelling.

A total of 12,929 people responded to the survey, of whom 4369 were Māori, 910 Pacific peoples and 1173 Asian peoples. These are total response ethnicity counts, where people who reported more than one ethnic group are counted in each group they reported.

Unlike the rest of the population where a sample survey was used, the Chatham and Pitt Islands over-sample was a census. Every dwelling was visited and every usually resident person that could be contacted was asked to participate.

#### Institutions survey

The institutions survey used a stratified sample design. Institutions were stratified according to four types. For each strata, establishments were randomly selected from lists provided by the Ministry of Health. These establishments were then visited and a small number of residents were randomly selected and interviewed from each. A total of 1076 residents responded to the survey.

## Summary of sample sizes

**Table 3:** Summary of sample sizes

| Survey | Total sample size | Ethnicity (total response) | | |
|---|---|---|---|---|
| | | Māori | Pacific | Asian |
| **Main survey (excluding Chatham and Pitt Islands over-sample)** | | | | |
| Design requirement | 12,000 | 4000 | 1000 | 1000 |
| Achieved | 12,529 | 4120 | 908 | 1172 |
| **Chatham and Pitt Islands over-sample** | | | | |
| Achieved | 400 | 249 | 2 | 1 |
| **Institutions survey** | | | | |
| Design | 1000 | | | |
| Achieved | 1076 | 97 | 13 | 6 |

## Main survey (excluding Chatham and Pitt Islands over-sample)

### Background

Originally, the proposed sample for the survey was 10,000 containing a minimum of 1500 Māori. After NRB won the tender, the sample was boosted to 11,000 and the stratification was to yield a minimum of 3000 Māori, 1000 Pacific peoples and 1000 Asian peoples. At this point no sampling had taken place.

After starting the sampling NRB was asked to boost the sample further by an additional 1000 Māori (ie, the total sample becoming 12,000). NRB continued with the targeting approach designed for the 11,000 sample. Concurrently, analysis was undertaken with respect to the expected design effects (DEFFs). Analysis showed inefficiencies in this design and a combination of targeting and screening was recommended instead.

Once this sample design change was agreed to, NRB stopped issuing more meshblocks and the recommended method was implemented.

### Sample design objectives

#### Sample size requirements

A requirement of the survey was a final sample size of at least 12,000. Within this total there were to be minimum subsamples of respondents belonging to following ethnic groups:

| | |
|---|---|
| Māori (total response) | 4000 |
| Pacific (total response) | 1000 |
| Asian (total response) | 1000 |

The expression 'total response' allows for respondents to nominate themselves as belonging to more than one ethnic group.

**Response rate requirements**

A requirement for this survey was a minimum true response rate of 70 percent.

## Expected sample sizes and accuracies

**Sample size**

To achieve the survey objectives, NRB worked out that 1200 PSUs or meshblocks would need to be sampled. NRB estimated that each PSU would yield an average of eight interviewed respondents, giving a minimum sample of 9600. In addition, a minimum of 2400 further people would be over-sampled by screening from a subset of the 1200 sampled meshblocks, by calling at each dwelling not already selected.

**Expected relative sampling errors**

After allowing for an estimate of the reduced efficiency of the sample arising from the stratification and clustering of the sample, NRB estimated that the following relative sampling errors (RSEs) would apply to the ethnic subsamples:

- Māori: up to 15 percent – within one of four equally sized age groups within gender.
- Pacific and Asian peoples: up to 20 percent – within one of two equally sized age groups within gender.

## Sample design – general comments

A stratified design was used. The strata were defined according to the ethnicity variable defined by Question 11 in the Individual Form in the 2001 Census.

The population density thresholds, which determine to which stratum a PSU belonged, were selected with the following points in mind:

- If the density threshold was set too high, there would be insufficient PSUs meeting that criterion to enable the final sample to yield the required ethnic subsample minimums.
- There was likely to be a direct relationship between increased density threshold and increased sample inefficiency. It was assumed that ethnic similarity within a PSU may well correspond with less variation in responses obtained for the variables being measured.

For some strata the PSUs could not be selected with probability proportional to size (in terms of the number of permanent private dwellings) due to the large number of PSUs required relative to the total number available within the strata. It would have resulted in larger PSUs having the possibility of being selected more than once. Hence, an equal probability of selection was chosen and offset by taking a fixed proportion of dwellings (ie, a constant sampling fraction) from each PSU to give equal probability of selection for all dwellings within a particular stratum.

Due to a change in the survey objectives shortly after the field work started the main survey has two sample designs.

### Initial sample design

The initial design consisted of four strata:

- Māori
- Asian
- Pacific peoples
- Other.

The Māori stratum consisted of all PSUs containing 60 percent or more eligible persons who identified themselves as Māori, according to the 2001 Census. In other words, the Māori stratum contained a high proportion of Māori people. PSUs were selected with equal probability. For each chosen PSU one out of two dwellings was selected. While 696 out of 861 PSUs were selected, only 32 were surveyed before the design changed.

The Asian stratum consisted of all PSUs containing 40 percent or more eligible persons who identified themselves as Asian. PSUs were selected with equal probability. For each chosen PSU two out of nine dwellings were selected. While 189 out of 340 PSUs were selected, only two were surveyed before the design changed.

The Pacific peoples stratum consisted of all PSUs containing 55 percent or more eligible persons who identified themselves as Pacific peoples. PSUs were selected with probability proportional to size. For each chosen PSU a constant sample of 12 dwellings was selected. While 125 out of 439 PSUs were selected, only three were surveyed before the design changed.

The Other stratum consisted of all remaining PSUs. PSUs were selected with probability proportional to size. For each chosen PSU a constant sample of 12 dwellings was selected. While 490 out of 36,712 PSUs were selected, only 77 were surveyed before the design changed.

### Latter sample design

The latter design consisted of two strata:

- Māori
- Other.

The Māori stratum consisted of all PSUs containing 70 percent or more eligible persons who identified themselves as Māori. In other words, the Māori stratum contained a high proportion of Māori people. PSUs were selected with equal probability. For each chosen PSU two out of three dwellings were selected. This sampling fraction was reduced in later stages of the survey to one in seven as a result of obtaining more Māori than expected. This change was referred to as “Māori modified”.

The Other stratum PSUs were selected with probability proportional to size. For each chosen PSU a constant sample of 12 dwellings was selected.

In addition to the two strata there were three ethnic over-samples within the Other stratum:

- Māori
- Pacific peoples
- Asian.

The sole purpose of the ethnic over-samples was to increase the number of respondents identifying themselves as Māori, Pacific peoples or Asian. Within these selected PSUs, all dwellings that had yet to be selected were screened. In other words all dwellings not already selected in the Other stratum were screened, resulting in some dwellings having more than one chance of selection.

For the Māori over-sample only respondents identifying themselves as Māori, Pacific or Asian were eligible for interview. The Māori over-sample consisted of all PSUs in the Other stratum containing 10 percent or more eligible respondents who identified themselves as Māori, according to the 2001 Census.

For the Asian over-sample only respondents identifying themselves as Asian or Pacific were eligible for interview. The Asian over-sample consisted of all PSUs in the Other stratum not already over-sampled and containing 30 percent or more eligible respondents who identified themselves as Asian, according to the 2001 Census.

For the Pacific peoples over-sample only respondents identifying themselves as Pacific were eligible for interview. The Pacific peoples over-sample consisted of all PSUs in the Other stratum not already over-sampled and containing 25 percent or more eligible respondents who identify themselves as Pacific peoples, according to the 2001 Census.

When the later design was implemented, NRB also over-sampled PSUs in the Other stratum who had been selected and surveyed in the initial design.

## Sample selection

### Primary sampling unit

The first stage of sampling took place at the meshblock level. A systematic sample was taken from each stratum starting from a random point. Every meshblock within each stratum had a known, non-zero probability of selection. No meshblock, either within a stratum or across the strata, had more than one chance of selection.

For some strata, meshblocks were selected with a probability proportional to their size in terms of the number of eligible persons; whereas, for other strata, meshblocks were selected with equal probability.

The following describes the method for selecting meshblocks from a stratum with a probability proportional to their size. For selecting meshblocks with equal probability of selection the method is equivalent if each meshblock is treated as having just one permanent, private dwelling.

The probability of selection for each meshblock is in direct proportion to the number of eligible dwellings within the meshblock. This unequal probability of selection was necessary to compensate for the fixed cluster size of interviews to be carried out within each meshblock. These two factors ensured equal probability of selection for every dwelling within the target population within each stratum.

For each stratum the meshblocks to be sampled were selected systematically, according to the following steps:

1. The list of meshblocks was sorted in order of District Health Board (DHB), but randomly within. This was to ensure a good representation of DHBs.
2. For each meshblock on the sorted list the cumulative number of permanent private dwellings was calculated according to 2001 Census estimates.
3. The skip (k) was chosen – the overall cumulative permanent private dwellings total was then divided by the number of PSUs that were to be drawn from it.
4. The initial sampling point was chosen – a random number (r) was generated with a value greater than zero but less than or equal to the skip (k).
5. The meshblock was selected by the one containing the rth private dwelling on the list being selected, then every kth unit down the list was selected. In other words the rth unit, r+kth unit, r+2kth unit etc were selected.
6. Continue skipping down the list until the required sample size was attained. This should be the point at which no more skips were possible.

### Secondary sampling unit

The secondary sampling unit is the dwelling. Each PSU is exactly described according to the streets, side of street and portion of the street that belong to the area.

Each PSU was assigned a sampling fraction. This sampling fraction was translated into an integer, which indicated the step between successive dwellings that were to be selected for inclusion in the survey. This integer was also designed to ensure complete coverage of the PSU.

Each selected PSU was visited and enumerated and the dwellings for inclusion in the survey were identified. This step ensured the number of eligible dwellings was updated from the census counts and enabled selection of a random starting dwelling.

### Respondent sampling

The final sampling stage was selecting the respondent. Within each dwelling all eligible people (ie, people those aged 15 years and over who lived at that dwelling) were identified. Within each selected dwelling one eligible person was selected. The Kish Grid was used to select the single eligible person from each selected dwelling. The names of all eligible respondents were listed in descending order of age onto the sampling grid. The respondent who was to be asked for an interview was the person whose name fell alongside a predetermined indicator.

### Chatham and Pitt Islands over-sample

The Chatham and Pitt Islands over-sample was a census of all permanent private dwellings visited. Every usually resident person who could be contacted was asked to participate.

A design requirement for this component of the survey was to achieve a minimum true response rate of 70 percent.

### Institutions survey

#### Sample design objectives

**Sample size requirements**

A requirement of the survey was a final achieved sample size of at least 1000 records. The decision was made on cost, rather than accuracy.

**Response rate requirements**

A requirement for this survey was a minimum true response rate of 70 percent.

#### Sample design

The stratified sample design used a list-based frame. The four strata were:

- hospitals
- IHC homes
- rest homes
- dependent persons.

Establishments were chosen and then a fixed sample size of respondents was randomly selected from within the establishments. To ensure a minimum sample size of 1000 a larger sample size was targeted to allow for non-response. More respondents were selected in the hospitals and rest homes strata as these establishments accounted for more residential beds in the population. Details of the sample design are recorded Table 4.

**Table 4:** Institutions survey – design details

| Strata | Required achieved respondent sample size | Estimated number of establishments selected* | Required targeted respondent sample size** | Achieved respondent sample size | Achieved response rate (%) |
|---|---|---|---|---|---|
| Hospitals | 300 | 100 | 429 | 340 | 92.6 |
| IHC homes | 200 | 67 | 286 | 209 | 88.9 |
| Rest homes | 300 | 100 | 429 | 322 | 89.0 |
| Dependent persons | 200 | 67 | 286 | 205 | 79.2 |
| Total | 1000 | 334 | 1430 | 1076 | 88.0 |

* Selecting an average three respondents per establishment (achieved).

** To allow for non-response of 30 percent.

### Sample selection for each stratum

The first stage of selection was choosing the establishments, the PSUs. Establishments were selected with probability proportional to their size, according to the number of residential beds. This resulted in larger establishments, in terms of the number of beds available, having a greater chance of selection. For each stratum the PSUs were sorted according to broad area, using telephone area codes.

Cumulative totals of residential beds were calculated, so each establishment was allocated a bed number range. A systematic sample was then taken, from a random start point, with the probability of selection for an establishment being proportional to size, as defined by the number of beds.

The sampling fraction used in the systematic procedure meant it was possible for establishments with numbers of beds greater than the sampling fraction to be selected more than once. This was permitted.

For most establishments respondents were selected using an alphabetical listing or bed number or room number system. When no such system existed, a random start point was chosen with the xth residents identified as the respondents.

Where residents were segmented into, for example, disability-type wards, a destratification of the wards and beds into an alphabetic or a number series needed to be undertaken before a random entry point was chosen and every xth person sampled.

Sampling anticipated a 70 percent response rate initially. However, participation rates proved to be higher resulting in the sampling fraction being altered during the course of the study, to prevent over-interviewing. Sample sizes of four, three and two respondents per establishment were applied as fieldwork progressed. Table 5 shows the number of establishments of each of the four types at which interviews were conducted, for each of the three cluster sizes, the yield and the response rates.

**Table 5:** Institutions survey – number of respondents selected per establishment

| Respondents selected per establishment | Number of establishments selected (unique) | Number of persons approached | Number of achieved interviews | Response rate (%) |
|---|---|---|---|---|
| **Clusters of four** | | | | |
| Hospitals | 75 (71) | 300 | 279 | 93.0 |
| IHC homes | 43 | 172 | 156 | 90.7 |
| Rest homes | 70 | 280 | 244 | 87.1 |
| Dependent persons | 58 | 232 | 180 | 77.6 |
| **Clusters of three** | | | | |
| Hospitals | 17 | 51 | 47 | 92.2 |
| IHC homes | 15 | 45 | 35 | 77.8 |
| Rest homes | 21 | 63 | 61 | 96.8 |
| Dependent persons | 9 | 27 | 25 | 92.6 |
| **Clusters of two** | | | | |
| Hospitals | 8 | 16 | 14 | 87.5 |
| IHC homes | 9 | 18 | 18 | 100.0 |
| Rest homes | 9 | 18 | 17 | 94.4 |
| Dependent persons | – | – | – | – |
| Total | | 1222 | 1076 | 88.1 |

Note: A large hospital could be selected more than once.

### Substitution

For both the main and institution surveys no substitution occurred in cases of non-response.

### Sample allocation

Ideally, samples should be allocated in such a way as to ensure that seasonal effect on health variables do not bias estimates produced from the survey. With the main survey more respondents were surveyed in the later half of the interview period, particularly more Asian peoples, Pacific peoples and Māori. Introducing a seasonality adjustment in the survey weighting was investigated, but seasonality was found to have no significant effect on any of the estimates examined. Nevertheless, care should be taken when analysing health data in which seasonality may be a factor.

## Questionnaire

The 2002/03 NZHS has four health-related and one demographic module. Most questions in the 2002/03 NZHS were drawn from health surveys used overseas or locally or had been compiled by researchers with expertise in the topics being addressed.

A copy of the 2002/03 NZHS questionnaire is available on the Ministry of Health's website: http://www.moh.govt.nz/phi

## Main questionnaire

### Chronic diseases

This module measured the self-reported prevalence of selected chronic diseases, including heart disease, stroke, diabetes, asthma, chronic obstructive pulmonary disease (COPD), arthritis, spinal disorders, osteoporosis and cancer. This module expanded the 1996/97 NZHS, which only included questions on diabetes and asthma.

Questions on chronic diseases were taken or adapted from previous New Zealand or international surveys on these topics after consultation with experts within and outside of the Ministry of Health. Almost all of the items are similar conceptually and in wording to those used in the 'Preliminary common instrument for chronic physical conditions' of the EUROHIS Project (Nosikov and Gudex 2003).

### Health service utilisation

This module measured the use of a wide range of health services, including general practitioners (GPs), Māori and Pacific health care providers, medical specialists, nurses, pharmacies and prescriptions, complementary and alternative medicine (CAM) providers, other health care providers (ie, physiotherapists, dieticians, dentists, opticians, social workers, occupational therapists, speech therapists and midwives), telephone or internet helplines, and public and private hospitals.

Questions covered frequency of contact, reasons for contact, satisfaction with services and reasons for any dissatisfaction, unmet health need, and reasons for unmet health need. This module expanded the 1996/97 NZHS, which focused on GPs, prescriptions and hospital use, and briefly covered other selected health care providers (ie, pharmacists, dentists, medical specialists, nurses, optometrists and physiotherapists).

Questions on use of GP and specialist services, use of other health services, prescriptions and over-the-counter medicines were taken or adapted from the 1996/97 NZHS (which had derived them from the 1992/93 Household Health Survey).

Questions on health services and providers not included in earlier surveys were developed in consultation with researchers with experience in these areas.

Additional items also had to be constructed for use of telephone and internet helplines. Some questions were modified so as to be consistent with those asked in the post-census Household Disability Surveys conducted in 1996 and 2001.

### Risk and protective factors

This module measured the prevalence of a range of health risk and protective factors, including high blood pressure, high cholesterol, overweight and obesity, adult weight gain, weight cycling, physical activity, tobacco smoking, marijuana smoking, vegetable and fruit intake, alcohol use and gambling. This module expanded the 1996/97 NZHS, which included only high blood pressure, smoking, physical activity and alcohol.

Questions about blood pressure, blood cholesterol, body weight, and vegetable and fruit intake were taken or adapted from previous New Zealand surveys, including the 1996/97 NZHS and 1997 National Nutrition Survey (Russell et al 1999). As with the chronic disease questions, they were largely compatible with the 'common instrument' being developed as part of the EUROHIS project (Nosikov and Gudex 2003).

The physical activity questions comprised the New Zealand Physical Activity Questionnaire Short Form, a local adaptation of the International Physical Activity Questionnaire that the Ministry of Health had developed in partnership with Sport and Recreation New Zealand and Statistics New Zealand and that the University of Auckland had validated (University of Auckland 2003). The physical activity questions differed from those used in the 1996/97 NZHS.

The tobacco smoking module was an adaptation of the World Health Organization (WHO) recommended instrument (WHO 1998). The alcohol module is the WHO Alcohol Use Disorders Identification Test (AUDIT) questionnaire (Babor et al 1992; Saunders et al 1993).

Questions on marijuana use were adapted from previous New Zealand drug surveys. Gambling questions were adapted from previous New Zealand and international gambling surveys.

### Self-reported health status

This module measured self-reported health status and covered general health, vision, hearing, digestion, breathing, pain, mental health, sleep, energy and vitality, understanding, communication, physical functioning, self-care, usual activities, and social functioning. The questions were derived from the SF-36 Health Status Questionnaire and the WHO Long Form questionnaire on health status. This section expanded the 1996/97 NZHS, which used only the SF-36 questionnaire. The reason for this change was because the SF-36 questionnaire did not reliably measure health status among Māori and Pacific peoples and because the WHO Long Form covered more health domains than the SF-36, so gives a better overall measure of self-reported health status.

Most new questions on health status were taken or adapted from the WHO Long Form instrument developed in association with the World Health Survey and based on the International Classification of Functioning, Disability and Health (WHO 2001). The SF-36 questions (version 1, Australia and New Zealand adaptation) were embedded in the WHO Long Form by slightly modifying, when necessary, the wording or response categories of the WHO Long Form items.

### Demographics

This module included standard sociodemographic variables such as age, gender, ethnicity and responses to ethnicity, country of birth, household characteristics, education, income support, employment, personal and household income, medical insurance and geographical area unit (for rural/urban and NZDep2001).

The demographic and socioeconomic questions were taken from the 2001 Census and/or the Household Labour Force Survey to ensure comparability with other major surveys. Minor variation in some questions (eg, ethnicity) was required to change from a self-completed to interviewer-administered format. The responses to ethnicity module derived some items from the United Kingdom Fourth National Survey of Minorities and others from the corresponding module of the United States Behavioral Risk Factor Surveillance System.

### Institutions and Chatham Islands questionnaires

The questionnaire used for the institutions survey was largely unchanged from the main survey.

The Chatham and Pitt Islands questionnaire was also largely unchanged. However, a few additional questions of local interest were added.

## Data collection

### Collection mode (method)

The collection mode chosen was face-to-face interviewing, using trained interviewers. The reasoning behind choosing face-to-face interviewing instead of telephone interviewing was:

- The 2001 Census showed that 3.7 percent of households in private occupied dwellings did not have access to a telephone.
- It has become accepted wisdom in the telecommunications industry, although not well quantified, that a proportion of homes no longer use landline phones, instead opting for a cell phone only approach (eg, younger households, single-person households and student flats).
- From an analysis that NRB commissioned on the Statistics New Zealand Household Income Survey (now the Household Economic Survey), non-ownership is markedly skewed to those Māori and Pacific peoples in the lower income groups (rather than those in the middle or higher), so the phone sample frame prejudices the people most in need of fair representation in the NZHS sample.

Using face-to-face interviewing can minimise this under-coverage, which has potential to bias the results. Face-to-face interviewing also ensures comparability with the previous health survey.

### Interview selection, training and performance

#### Selection of interviewers

Selection of competent interviewers was a key step to obtaining a good response rate. NRB analyses every survey it conducts in terms of the individual response rate achieved by each interviewer. Factoring in the urban/rural dimension, interviewers are ranked on their ability to achieve responses. Not all interviewer applicants will, even after specific training, achieve the necessary response rate. Selection tools by way of attitude and cognitive tests help reduce the number, but like most psychometric tests are only a modest help in field performance prediction. This means interviewers need to be screened, trained and monitored, and then exited or removed to less demanding commercial or phone assignments, where 50 percent is considered acceptable by the market. Replacement, new training and "recovery" sweeps over the exited person's work

are required. This select, train, trial, emit, reselect and recovery activity is costly, but produces high response rates.

For the Chatham and Pitt Islands over-sample, interviewers were drawn from the mainland to avoid local staff inevitably interviewing people they knew about possibly sensitive health matters.

### Training and in-field support

Interviewers need to have a strong sense of 'entitlement' to approach homes, a strong sense of the 'value of the survey' to sell participation to the household, and a versatile selection of 'engagement options' with which to ensure they can find an intercept between the eligible person's time/place/attitude configuration, and the time required for the interview.

In-field support included assistance with the names of occupants of persistently empty homes. The electronic electoral role was searched for the name, then the telephone book for the telephone number, and then the person was contacted for an interview time. This was then fed back to the interviewer to implement. Similar strategies were used to contain dogs in some areas and to access apartment blocks in built-up areas.

Interviewers and field supervisors received a two-day training course on how to conduct interviews. Material used in the training included:

- a meshblock map for household enumeration and selection
- instructions for sampling dwellings
- a letter from the Ministry of Health
- a brochure describing the survey
- a Kish Grid for respondent selection
- a consent form
- the questionnaire
- show cards
- scales for weight and a portable stadiometer for height
- two tests to assess trainees' pick-up of training
- thank you cards.

Field supervisors also received additional training on:

- contact and support with interviewers
- progress and evaluation forms for interviewers.

### Performance

The interviewers' performance was regularly monitored. They were rewarded for *applying* their training successfully.

## Call pattern

In addition to good interviewer preparation the call pattern is an important component of achieving higher response performance.

### Number of calls

The 'call' refers to one visit on one day during a particular time band, eg, 5–8 pm. While a second or third attempt might be made to a given dwelling while the interviewer was in the area, these were attempts and not calls in the current definition. The reason for this distinction was that the second and subsequent attempt on a given day was less productive. NRB conducted a total of up to 10 calls at each sampled dwelling before accepting that dwelling as a non-contact dwelling. (Ten calls might entail up to 20 or more attempts.)

### Spacing of callbacks

Making all 10 callbacks within a fortnight does not capture people who are away from their dwelling for a fortnight or longer. Therefore, the procedure was to make six calls in the survey month that the meshblock was issued, pause for three to four weeks, attempt two more calls, and finally pause a further three to four weeks before implementing the final two callbacks. This helped not only in the case of people temporarily away, but also with people who were in a burst of activity work-wise, socially or for family reasons when their dwelling was first approached.

The interviewing team made two separate visits to the Chatham and Pitt Islands. The purpose of the second visit was to do additional callbacks at homes 'too busy', vacant or 'unavailable' at the first visit, and to attempt refusal recovery. The first sweep yielded 348 interviews and the second 52 interviews, a total of 400.

### Recovery attempts

When a meshblock or interviewer-meshblock combination produced a below standard response rate, it was identified in the field manager's computer tracking of the response rate. After investigation, a different interviewer might have been assigned to re-visit that meshblock to re-attempt the refusals, generally with an explanation to refusers why an interviewer was approaching them again.

### Invigilation

This step refers to the field supervisor in each area phoning back the household to confirm the interview was done and check the respondent was the one stated. It is rare for interviewers who are properly trained and field-supported to falsify interviews, but nevertheless NRB perform a 15 percent field check on each interviewer as a precautionary measure.

### Enumeration and pre-survey letter

In implementing a meshblock sample, two steps were helpful in improving the capture of the interviews. The first was to re-enumerate the meshblocks to take into account the number of new dwellings built, or in some cases the number of buildings demolished, over the period since the last pre-census enumeration. This was a period of about two years. The second was to drop a pre-survey letter at those dwellings that fell on the 'every-xth' dwelling sampling fraction. An advance letter has been shown to improve response rates in some cases, although the improvement is more marked for phone interviewing than face-to-face interviewing.

### Consent

The survey was voluntary. Adults who were selected were told about the survey and given an information brochure. If they agreed to take part, they were asked to sign a consent form. Information brochures and consent forms were available in a range of languages other than English. It was also possible to match respondents and interviewers by language, ethnicity and sex.

### Proxy reporting

With the exception of the institutions survey no person answered questions on behalf of another person. For the institutions survey caregivers were able to provide responses on behalf of the person for whom the interview was taking place. Of the 1076 responses, 734 (68.2%) were proxies (521 required substantial assistance and 213 required some assistance). Cases where proxy reporting was used were recorded and flagged in the final dataset.

### Field dates

The pilot test (sample size, n=114) was conducted from August to September 2002.

The main survey was undertaken from September 2002 to January 2004, with the Chatham and Pitt Islands over-sample (n=400) completed from September 2002 to November 2002.

The institutions survey (n=1076) was undertaken from December 2002 to May 2003.

## Field testing

### Cognitive pilot test

This was a test of the acceptability and understandability of the questions carried out by NRB in July 2002. Respondents were 12 focus groups covering a range of ages, both genders and four ethnic groups.

In summary, none of the focus groups experienced much difficulty with any of the questions in terms of meaning, comprehensibility or sensitivity. However, the wording of a few questions and some interviewer instructions and show cards were slightly amended.

### Conventional pilot test and dress rehearsal

This was a test of the performance of the questionnaire and the survey design as a whole. NRB completed the pilot (sample size = 114) from August to September 2002 and achieved a 72 percent response. Although naturally small in scale, the pilot engaged the full range of survey materials and processes proposed for the national rollout. This extended all the way to the production of a unit record dataset and a data dictionary, and a small selection of tables for proofing purposes.

The objectives of the pilot test were to:

- check the response rate to participation in a one-hour long interview with particular reference to the key performance requirement of 70 percent
- check the participation in the more sensitive questions in the questionnaire, in particular marijuana use, weight measurement and personal income
- determine whether every question, disregarding choice of words, had a code for every response option or a provision for other responses
- examine the effectiveness of the Kish Grid in its doorstep implementation by interviewers for delivering age and gender spreads approximating the New Zealand population
- capture the data from a sample of more than 100 interviews to ensure every answer option could be scanned or punched into the electronic dataset, and to prepare a selection of percentage tables enabling examination of the completeness of the data when expressed in its final end-use format
- develop a final dataset layout.

#### Testing region

The pilot was implemented in the Wellington Regional Authority area. The reason for using this area was that the Wellington Ethics Committee was able to give approval for the pilot within its jurisdiction, whereas formal signoff from the other ethics committees was pending. This region also provided a good composition of persons from different ethnic and socioeconomic groups.

#### Meshblock selection

To reflect a reasonable comparison with the national field task, the 15 meshblocks were selected to reflect variation in ethnicity and socioeconomic status:

- three were chosen with a Pacific peoples emphasis
- two were chosen with an Asian peoples emphasis
- 10 were chosen with a socioeconomic emphasis.

#### Respondent selection

The Kish Grid was used as a means of selecting one person to interview from all of those eligible in the household. The Kish Grid is a method of selecting a person, free from interviewer bias, and free from 'door opener' bias or biases generated by the availability or preferences of the occupants.

The performance of the Kish Grid was monitored, particularly gender balances and cultural effects. As a result of the field test, enhancements were made to training, monitoring and procedures.

#### Interviewer training

Seven interviewers were trained over two days to conduct the pilot interviews. Three field supervisors received the same training plus additional training.

#### Questionnaire

Most questions asked were drawn from health surveys used overseas or locally or had been compiled by researchers who specialised in the topics being addressed. Given this background, it was not expected that many question changes would emerge from the field pilot.

Refusal of specific or sensitive questions within the questionnaire was very low. A few refusals were received for some of the sensitive questions. NRB's quality controls traced these to an interviewer effect. Procedures were taken to ensure no repeat of such effects in the main survey. As a result of the field trial a few changes were made to the questionnaire.

#### Response rate

The key performance indicator for the survey was a true response rate of at least 70 percent. This response rate was achieved, notwithstanding the shorter callback period due to containing the time-frame for the pilot. The response rate at the cut-off was 72 percent. The response rate was monitored and analysed by meshblock and interviewer.

#### Data processing

The data from the pilot interviews were electronically captured, edited and formatted. A few unweighted tables were produced. No weighting estimation was performed for the field pilot.

## Response rates and respondent load

### Response rates

Not only was the survey well received by the New Zealand public, but most of the respondents agreed to be contacted again in two to three years:

- 72 percent of eligible people approached for the main survey (excluding the Chatham and Pitt Islands over-sample) completed a questionnaire, with 88 percent of these respondents agreeing to be contacted again in two to three years
- 73 percent of eligible people approached for the Chatham and Pitt Islands over-sample completed a questionnaire, with 99 percent of these respondents agreeing to be contacted again in two to three years
- 88 percent of eligible people approached for the institutions survey completed a questionnaire, with 81 percent of these respondents agreeing to be contacted again in two to three years.

There were essentially four components to the response rate calculation:

1. Ineligibles (eg, vacant sections, vacant dwellings and non-residential dwellings).
2. Eligible responding.
3. Eligible non-responding.
4. Unknown eligibility (eg, non-contacts and refusals who provide insufficient information to determine eligibility).

The response rate was 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 this response rate was that a proportion of the unknowns were likely to be eligible if contact could have been made. As we were not able to make contact with the estimated number who would be eligible, they were classified as non-respondents.

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

The official response rate measure used unweighted counts, which are useful for monitoring field operations. A minimum response rate of 70 percent was a design requirement of the survey. The response rate was monitored and analysed by PSU and by interviewer throughout the survey collection period.

Unweighted response rates, computed using the number of people selected, reflect the success of the operational aspects of the study (ie, getting the selected people to participate). Response rates weighted to reflect the probability of being selected into the sample describe the success of the study in terms of the population being measured.

The unweighted and weighted response rates will be the same when every person selected for the survey has the same chance of selection. However, if people are selected with different chances of selection, there will be a difference in the response rates.

Future iterations of the survey will use weighted counts as the official measure.

### Respondent load (burden)

The high response rate and high proportion of people agreeing to be contacted again in two to three years indicate that the survey was well received by the public.

For the main survey, the median time taken by participants to complete the interview was 60 minutes, with the lower and upper quartiles being 46 and 70 minutes and the longest, 205 minutes.

For the institutions survey, the median time taken to complete the interview was 50 minutes, with the lower and upper quartiles being 40 and 65 minutes and the longest, 205 minutes.

### Measures used to maximise response and minimise respondent load

The survey and processes were carefully designed to ensure the impact on respondents was minimised. The following measures were used to maximise the response rate.

- Only one eligible person was selected per dwelling.
- A well-tested and largely well-proven questionnaire was used.
- Skilled interviewers carried out the interviews.
- Appointments were taken for interviews.
- Interviews were accepted away from the dwelling in special circumstances.
- Language, ethnicity and sex matching of eligible respondent and interviewer were used where necessary.
- Well-designed call pattern processes were used.
- Interviewers had in-field support.
- Interviewers were monitored and, if found to be underperforming, were removed or replaced if necessary.

## Data processing

NRB was responsible for data capture, editing and coding.

### Data capture

The data collected via a pen and paper questionnaire was captured electronically using systems tested in the pilot test.

### Editing

The unit record datasets provided to the Ministry of Health were edited for range and logic. Any inconsistencies found have been remedied by returning to the questionnaire and, if necessary, to the respondent for clarification and correction.

### Imputation

No explicit unit record or item imputation was used in the survey to deal with unit record or item non-response. However, non-response has been implicitly adjusted for in the weighting estimation by benchmarking the survey population to an estimate of the target population.

All respondents who had covered all the questions (regardless of whether each response was don’t know, not specified or refused) and who had provided their sex and year born were included in the weighting estimation. Otherwise, they were excluded from the weighting estimation and treated as non-respondents.

### Coding

Most of the questions used a single tick box, although some questions:

- Asked for the person's age when a disease was diagnosed or for the number of visits to a particular health provider.
- Offered an 'other' category where respondents could specify non-standard responses. Each other category response was recorded and later analysed. Each response was categorised to an existing code, coded to a newly set up 'standard' code or coded as other.
- Allowed multiple responses. For these questions, all responses were retained with each response shown as a separate variable on the data file (ie, Q12_01, Q12_02, Q12_03, etc).

### Dataset delivery

NRB was contracted to deliver the following interim and final datasets and accompanying documentation (data dictionary) to the Ministry of Health:

- pilot
- main survey (interim datasets every 2000 records)
- Chatham and Pitt Islands
- institutions.

The Ministry of Health received the final datasets and documentation for the main survey, institutions survey, and Chatham and Pitt Islands in February 2004.

## Quality control

Quality control of data and processes was an integral component of the survey implementation. It was implemented through comprehensive testing, ongoing performance monitoring, peer review, using standard classifications and concepts (where possible), and using specialist staff.

### Interviewers

Quality control of interviewers meant interviewers:

- were selected after taking a written test and having a personal interview to screen for above average aptitudes
- received explicit stepwise training in each identified task and risk element
- were tested after training to check on their uptake and retention
- were monitored in the field and, when necessary, were removed and replaced if under-performing.

### Interviewing

Quality control of the interviewing meant:

- field checking of 15 percent of completed interviews to ensure fidelity

- an area supervisor monitoring and counselling on in-the-field issues day to day
- support to the field interviewers from a trouble-shooter at head office to raise the field success rate
- recovering interview reattempts and refusals.

### Call pattern

Quality control of the call pattern meant:

- undertaking callbacks at different times of the day or different days of the week
- separating the first and subsequent visits to maximise the opportunity to contact the occupants
- taking appointments for interviews
- accepting interviews away from the dwelling in special circumstances to maintain sample mix.

### Field work allocation

Quality control of the field work allocation meant:

- visually matching the meshblock label against the interviewer's recorded meshblock to ensure they were correct to eliminate transcription errors
- visually matching the interview number against the sample sheet number to ensure the household was correctly numbered
- checking the address of each interview to ensure each interviewer had not strayed outside their given meshblock.

### Outcomes

Quality control of the outcomes meant:

- continuously monitoring the response rate per meshblock
- matching the language, culture or gender of an eligible respondent and an interviewer to minimise non-response bias.

### Editing

Quality control of the editing meant:

- field editing by area supervisors to ensure completion
- supervisors re-contacting respondents if data was missing data
- an electronic edit to determine whether there were duplicate serial numbers or non-eligible serial numbers, reverting to paper copies if an error was found
- running an electronic check over the branching 'skip to' instructions throughout the questionnaire to identify overfills for removal and underfills for return to the field for follow up (generally by phone)
- running an electronic range check to ensure all data fell within the permissible code range, then checking the questionnaire or re-contacting the respondent if necessary

- checking the range and logic to identify inconsistencies and errors, then checking the questionnaire or re-contacting the respondent if necessary
- checking interview start and finish times to correct errors made by interviewers using the 24-hour clock.

### Dataset and documentation

Quality control of the dataset and documentation meant:

- creating a selection of tables to check that row and column totals balanced and were correct compared with total records captured, including examining cell counts in cross tabulation to ensure integrity of the tables
- carrying out independent pen and paper calculations on a selection of interviews for derived variables to check against SAS program creation of these same scores
- creating a data dictionary including a detailed description of each variable and response value.

## Weighting estimation

The Ministry of Health carried out the weighting estimation once quality checks were completed.

### Main survey (excluding Chatham and Pitt Islands over-sample)

#### Selection weight

As the survey was conducted on a sample of respondents, each person represented a number of other people in the population. Therefore, each respondent had a weight that indicated how many population units were represented by the sample unit.

Survey weights allow a sample to be used to produce estimates for the entire population. The selection weight (SELWGT) is the sample design weight before any adjustment for unit record non-response and post-stratification.

The calculation of the selection weight was complex. The methodology was complicated by having a 'main' sample and an over-sample component. The following is a general description of the methodology.

For each stratum, the probability of selecting a person for the 'main' sample = W1 x W2 x W3.

Where:

W1 = (total number of chosen 'main' sample dwellings in the stratum/total number of dwellings in the stratum). This is the probability of selecting the PSU, but using dwelling counts instead of PSUs to account for the PSUs not being of equal size in terms of dwellings.

W2 = (number of chosen dwellings in the selected PSU/total number of dwellings in the selected PSU). This is the probability of choosing a dwelling within a selected PSU.

W3 = **(number of chosen eligible persons in the selected dwelling/total number of eligible persons in the selected dwelling).** This is the probability of choosing an eligible person within a selected dwelling.

However, due to over-sampling within some PSUs the calculation gets more complex, as some respondents have more than one chance of selection.

As noted in the sample design section the following over-samples existed:

- Māori (the eligible were Māori or Pacific or Asian peoples)
- Asian (the eligible were Pacific or Asian peoples)
- Pacific (the eligible were Pacific peoples).

For PSUs where over-sampling occurred the above eligible persons had two chances of selection: selection for the 'main' sample and selection for the over-sample.

For PSUs where over-sampling occurred the probabilities of selection were:

$$\text{Probability of selection} = \left[\begin{array}{l}\text{probability of selecting a}\\ \text{respondent for the 'main' sample}\end{array}\right] + \left[\begin{array}{l}\text{probability of selecting a respondent}\\ \text{from screened dwellings}\end{array}\right]$$

$$= (W1 \times W2 \times W3) + (W1 \times W2_s \times W3_s)$$

Where, M = main sample and S = over-sample (screened).

For example, PSUs selected where Asian over-sampling occurred the probability of selection was:

- For Asian or Pacific respondents, regardless of whether they were selected from a screened or non-screened dwelling:
  Probability of selection $= (W1 \times W2 \times W3) + (W1 \times W2_s \times W3_s)$
- For non-Asian and non-Pacific respondents:
  Probability of selection $= (W1 \times W2 \times W3) + 0$
  $= (W1 \times W2_s \times W3_s)$

Similarly, for PSUs where other over-sampling occurred.

The inverse of the probability of selection is the survey weight. In other words, the selection weight is one divided by the probability of selection.

Due to a change in the survey objectives shortly after the field work started two sample designs were implemented. Several PSUs selected in the initial design were surveyed before the change in sample design. These PSUs had no chance of selection in the later design.

The samples from each design were weighted so that as a whole they represented the New Zealand population, with the original design's selection weights preserved. The selection weights for the later design were calculated so they accounted for that part of the population not covered by the initial design.

### Non-response adjustment

No explicit non-response adjustment was performed.

### Final weighting

The final stage of the weighting process is a weighting adjustment using generalised regression to ensure the final weighted totals of eligible adult respondents are consistent with independent population estimates. The 2002/03 NZHS was benchmarked to the 2001 Census population.

This adjustment also adjusts for under-coverage in the frame and non-response and reduces the level of sampling error for benchmark variables.

Generalised regression weighting is an improvement on more traditional post-stratification weighting, by allowing the flexibility to incorporate several population benchmarks. The initial selection weights were modified to produce final survey weights that aggregate to independent population estimates (or benchmarks).

The following benchmarks were used for the number of people aged 15 and over living in permanent private dwellings:

- by sex by 10-year age group (15–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75+)
- by sex by prioritised ethnicity (Māori, Pacific peoples, Asian peoples and European/Other) by lifecycle age groups (15–24, 25–44, 45–64, 65+)
- by sex by NZDep2001 deciles.

### Chatham and Pitt Islands over-sample

As this was a census all respondents had equal probability of selection with a selection weight of one. However, not all eligible persons responded to the survey. Therefore, a post-stratification adjustment was made to ensure the final weighted totals of eligible adult respondents were consistent with independent population estimates from the 2001 Census.

The post-stratification factors were calculated for sex by ethnicity (Māori, non-Māori) by age group (15–24, 25–34, 35–44, 45–54, 55–64, 65+).

Once the weighting for the main survey (excluding the Chatham and Pitt Islands) and the Chatham and Pitt Islands over-sample components were weighted the resulting datasets were joined.

### Institutions survey

As the institutions survey was conducted on a sample of respondents, survey weights were calculated for each respondent to the survey.

For each of the four strata, the probability of selecting a person = W1 × W2.

Where:

W1 = (total number of chosen residential beds within the stratum/total number of residential beds within the stratum).

As establishments were selected with probability proportional to size, then:
= ((number of PSUs chosen) x (total number of residential beds within the selected PSU))/(total number of residential beds within the stratum).

This is the probability of selecting the PSU, but using residential bed counts instead of PSUs to account for PSUs not being of equal size.

W2 = (number of residential beds chosen within the selected PSU/total number of residential beds within the selected PSU). This is the probability of choosing a person within a selected establishment.

Note:

W1 × W2 = ((number of PSUs chosen) x (number of residential beds chosen within the selected PSU))/(total number of residential beds within the stratum)

The inverse of the probability of selection is the survey weight. In other words, the selection weight is one divided by the probability of selection.

### Age-standardised weights

Age is an important determinant of health status. Therefore, when making comparisons between males and females, and between different ethnic groups, the different age distribution of the comparison population must be taken into account. Age-standardisation was performed by the direct method using the WHO world population age distributions applied to population counts from the 2001 Census. In addition to the New Zealand population survey weight an age-standardised weight exists.

For each prioritised ethnic group (Māori, Pacific peoples, Asian peoples and Other) and sex, the New Zealand population age distributions were adjusted to match the WHO world standard population distribution. Therefore, each prioritised ethnic group was treated as having the same age structure for each sex. By using age-standardised weights any differences between ethnic groups for either sex, cannot be attributed to differences in the age structure, as this factor has been accounted for.

### Replicate survey weights

The unit record data contains replicate survey weights to enable confidence intervals to be calculated easily.

## Data reliability

Two types of error are possible in an estimate based on a sample survey: sampling error and non-sampling error.

Estimates from this survey are subject to sampling errors or variability because they are based on information relating to a sample of persons rather than a full enumeration. That is, they may

differ from the estimates that would have been produced if the information had been obtained for all people. The method for estimating sampling errors is outlined below.

Other inaccuracies can occur because of insufficient coverage of respondents, inadequacies and imperfections in answers provided by respondents, and errors made when coding and processing data. Such inaccuracies are referred to as non-sampling errors and may occur in any enumeration regardless of whether it is sample or full enumeration. Significant efforts have been made to reduce non-sampling errors by carefully designing and testing the survey, questionnaire and processes, and ensuring detailed quality control of procedures and data.

## Sampling errors

### Main method

The Ministry of Health calculated sampling errors for survey estimates using a replicated method, called the Delete-a-Group (DAG) jackknife method (Kott 1998). The idea behind the replication approach was to divide the sample into G random groups, and then estimate the variance of full sample survey estimate. For the 2002/03 NZHS, 100 random groups were chosen (G=100).

#### Calculation of the replicate weights

G subsamples were produced by deleting one group at a time from the full sample. Each member of the full sample was assigned to a group in a way that mirrored the sample design. This was done so each subsample replicated the design of the full sample, but contained slightly fewer members. Each subsample was then reweighted to the population using exactly the same weighting estimation methodology as for the full sample.

Therefore, each record on the unit record dataset has G 'replicate weights' associated with it in addition to the 'main weight'. The replicate weights enable calculation of the confidence interval on each estimate.

#### Calculation of the confidence interval using the replicate weights

For each estimate another G replicate estimate was calculated using the G replicate weights. The variance of the full sample statistic was estimated using the variability among the G replicate estimates. This was done by taking the sum of the squared differences between the G replicate estimates and the original full sample estimate, and multiplying this by (G-1)/G.

To summarise, the formulae for calculating the variance of an estimate using this method are:

$$\text{variance}(y) = \frac{(G-1)}{G} \times \sum_g (y_g - y)^2$$

where

$G$ = 100 (the number of replicate groups)
$g$ = 1, 2, ..., G
$y_g$ = weighted estimate, having applied the weights for replicate group g
$y$ = weighted estimate from the full sample.

For the 95 percent confidence interval:

Sampling error ($y$) = $1.96 * \sqrt{\text{variance}(y)}$

Confidence interval ($y$) = $y \pm$ sampling error ($y$).

The near unbiasedness of the DAG jackknife requires the number of first-phase samples in each stratum to be large, say, greater than five.

The sampling errors have been represented as 95 percent confidence intervals. The 95 percent confidence interval of a survey estimate provides an indication of the margin of sampling error for that estimate. We are 95 percent certain that the true estimate lies within this range.

### Method for quantiles

An indirect method known as the Woodruff method has been used for quantiles. See Sarndal et al (1992) for details about this method, which is a transformation process, incorporating the DAG jackknife approach.

The Woodruff method indirectly estimates the variance of a quantile by calculating a confidence interval on the quantile from a cumulative density function (CDF) and then using the width of the confidence interval to derive a variance estimate. Thus, the variance is not estimated directly from the variation of the replicate estimates.

The Woodruff method was used because it was not clear that the DAG jackknife method would provide reliable estimates for the variance of quantiles, such as medians (Rao et al 1992).

## Classifications and standards

Standard classifications have been used where appropriate to promote comparability and data consistency.

## Security of information

Any information collected in the survey that could be used to identify individuals has been treated as confidential. Names and addresses of people and households collected in the survey have not been stored with their responses. No information will be released in a way that would enable an individual or a household to be identified.

Unit record data is stored in a secure area and is accessible on a restricted 'needs to know' basis only. All applications by academics or researchers to access anonymised unit record files will be assessed according to predefined criteria. If successful, applicants will be required to sign an agreement to ensure no breach of confidentiality occurs in regards to the storage of, and access to, the data and their outputs.

## Liability

Care and diligence has been taken to ensure the information in this document is accurate and up to date. However, the Ministry of Health accepts no liability for the accuracy of the information, its use or the reliance placed on it.

## Dissemination of data

The Ministry of Health has planned a series of outputs based on the 2002/03 NZHS findings. Where appropriate, collaboration will occur within the Ministry of Health and with external experts.

The first publication, *A Snapshot of Health: Provisional results of the 2002/03 New Zealand Health Survey*, was released in December 2003 (Ministry of Health 2003). This report was based on provisional data (ie, an incomplete dataset and an interim survey weighting estimation) and included selected key results stratified by gender and ethnicity.

The second publication is this descriptive report and datacubes containing additional descriptive results. This report comprises five chapters: introduction and methodology, chronic diseases, risk and protective factors, health service utilisation, and self-reported health status. The report is based on data from 12,929 adults living in private houses in New Zealand (including the Chatham and Pitt Islands).

Other planned outputs include estimates for DHBs, comparisons with previous surveys, results for the separate institutions survey, additional self-reported health status results, and results from more in-depth analyses examining interrelationships between the various health modules (eg, chronic diseases will be examined in relation to the prevalence of biological and behavioural risk factors for these diseases). All future outputs and information about how to obtain these outputs will be available on the Ministry of Health's website: **http://www.moh.govt.nz/phi**

### This report

The purpose of this report was to disseminate survey results as quickly as possible in an easily accessible format. Therefore, this report is purely descriptive, with sufficient commentary to facilitate understanding of the data. Even with analyses restricted to key sociodemographic indicators (ie, sex, age, ethnicity and NZDep2001), not all results could be included in the report. However, additional results are included in the datacubes.

It is hoped that this report, together with the datacubes, will meet the needs of most people and organisations. For those who are interested in more detailed information, the various methods for obtaining this, including how to access the dataset, are outlined on the Ministry of Health's website: http://www.moh.govt.nz/phi

#### Data analysis

##### Survey weights

Survey weights were applied to all analyses to produce nationally representative estimates. Survey weights were unique to each respondent and adjusted for individual probability of selection and differential non-response. The method used to calculate survey weights is described in the methodology section.

## Data reliability

Ninety-five percent confidence intervals were used to represent sampling error. The 95 percent confidence interval of the survey estimate provides an indication of the margin of sampling error for that estimate. The method used to estimate sampling error is described in the methodology section.

Ninety-five percent confidence intervals are presented for all descriptive results, either following the estimate in the text or summary tables at the end of the chapter, or as error bars in graphs. In this report, when a difference between population subgroups is referred to as significant, it means the difference is statistically significant at the 95 percent confidence level (ie, the 95 percent confidence intervals do not overlap).

When an unweighted individual cell contained a value of less than 10, results were suppressed for reasons of reliability and confidentiality.

## Age-standardisation

Age is an important determinant of health status and health risks. Therefore, when making comparisons between population subgroups (eg, between sex or ethnic groups), the different age distribution of the comparison populations must be taken into account.

A process called age-standardisation was used to adjust for the differing age distributions of population subgroups. Age-standardisation was performed by the direct method using the WHO World Population as the standard population. The methods used for age-standardisation are described in the methodology section.

## Subgroup analyses

An objective of the survey was to examine differences between population subgroups (ie, sex, ethnicity, age and NZDep2001).

### Sex

Most results were stratified by sex. All sex-specific estimates presented in the body of the report are age-standardised. Crude sex-specific estimates are in summary tables at the end of each chapter or in the datacubes.

### Age

Age was calculated using information on year of birth and date of interview. Where possible, age-specific estimates were calculated for the following 10-year age groups: 15–24, 25–34, 35–44, 45–54, 55–64, 65–74 and 75+. Otherwise the following lifecycle age group were used: 15–24, 25–44, 45–64 and 65+.

### Ethnicity

The classification of ethnic group was based on respondents indicating with which ethnic group or groups they identified. When only one ethnic group was given, that group was assigned. When two or more ethnic groups were given, a single ethnic group was assigned using the following priority rules:

- If Māori was one of the groups reported, the respondent was assigned to 'Māori'.
- If any Pacific ethnic group was reported, the respondent was assigned to 'Pacific'.
- If any Asian ethnic group was reported, the respondent was assigned to 'Asian'.
- All other respondents were assigned to 'European/Other'.

When possible, ethnic-specific estimates were calculated for Māori, Pacific, Asian and European/Other ethnic groups. When Pacific and Asian ethnic groups were not represented in adequate numbers for reliable estimates they were combined with European/Other to give a non-Māori ethnic group or suppressed. All ethnic-specific estimates included in the body of this report are age-standardised. Crude ethnic-specific estimates are in the summary tables at the end of each chapter or in the datacubes.

### Deprivation

NZDep2001 was used as the key indicator of socioeconomic status. It is an area-based index of deprivation that measures the level of deprivation for each meshblock, according to a combination of Census 2001 variables (ie, income, transport (access to car), living space, home ownership, employment status, qualifications, support (sole-parent families) and access to a telephone).

There are 10 NZDep2001 categories (deciles), with decile 1 representing the least deprived 10 percent of small areas and decile 10 representing the most deprived 10 percent of small areas. For this report, NZDep2001 deciles were aggregated into quintiles 1–5, with quintile 1 being the least deprived quintile and quintile 5 being the most deprived. All results presented by NZDep2001 quintile in this report are age-standardised.

### Interpretation of results

All results presented by sex, ethnicity and NZDep2001 in the body of this report have been age-standardised. This is to allow comparisons between population subgroups without differences in the age distribution of the comparison populations influencing results. Therefore, to compare one population with another, use the age-standardised estimates. However, note that age-standardised estimates have no meaning by themselves; they are meaningful only when compared with other age-standardised estimates.

If you want to know the actual burden experienced by the population of interest (eg, the prevalence of obesity or diabetes in males or females), use the crude results shown in the summary tables at the end of the relevant chapter or in the datacubes.

# References

Babor TF, de la Fuente JR, Saunders J, et al. 1992. *AUDIT. The Alcohol Use Disorder Identification Test. Guidelines for Primary Care.* Geneva: World Health Organization.

Kott PS. 1998. *Using the Delete-A-Group Jackknife Variance Estimator in NASS Surveys.* RD-98-01. Washington, DC: USDA.

Ministry of Health. 1999. *Taking the Pulse: The 1996/97 New Zealand Health Survey.* Wellington: Ministry of Health.

Ministry of Health. 2002. *The New Zealand Health Monitor.* Public Health Intelligence Occasional Bulletin No 13. Wellington: Ministry of Health.

Ministry of Health. 2003. *A Snapshot of Health: Provisional results of the 2002/03 New Zealand Health Survey.* Wellington: Ministry of Health.

Nosikov A, Gudex C (eds). 2003. EUROHIS: Developing common instruments for health surveys. *Biomedical and Health Research* 57. Amsterdam: IOS Press.

Rao JNK, Wu CFJ, Yue K. 1992. Some recent work on resampling methods for complex survey. *Survey Methodology* 18: 209–17.

Russell DG, Parnell WR, Wilson NC, et al. 1999. *NZ Food: NZ People: Key results of the 1997 National Nutrition Survey.* Wellington: Ministry of Health.

Sarndal CE, Swensson B, Wretman J. 1992. *Model Assisted Survey Sampling.* New York: Springer.

Saunders JB, Aasland OG, Babor TF, et al. 1993. Development of the Alcohol Use Disorders Screening Test (AUDIT). WHO collaborative project on early detection of persons with harmful alcohol consumption. II. *Addiction* 88: 791–804.

Statistics New Zealand, Ministry of Health. 1993. *A Picture of Health.* Wellington: Statistics New Zealand and Ministry of Health.

University of Auckland. 2003. Validation of MOH-short and SPARC-long physical activity questionnaires. Unpublished report to the Ministry of Health and Sport and Recreation New Zealand.

WHO. 1998. *Guidelines for Controlling and Monitoring the Tobacco Epidemic.* Geneva: World Health Organization.

WHO. 2001. *International Classification of Functioning, Disability and Health.* Geneva: World Health Organization.