2014 Health and Lifestyles Survey

Methodology Report

health promotion agency

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

1. INTRODUCTION

The Health and Lifestyles Survey (HLS) is a biennial monitor of the health behaviour and attitudes of New Zealand adults aged 15 years and over, and parents and caregivers of 5 to16-year-olds, first carried out in 2008. The HLS is managed by the Health Promotion Agency (HPA)[1] and collects information relating to HPA's programme areas of alcohol, tobacco control, sun safety, minimising gambling harm, nutrition and physical activity, mental health and immunisation.

The 2014 HLS involved face-to-face interviews with 2,594 adults (aged 15 years and over). Some of these adults (n=463) were also included in interviews of 742 parents and caregivers of 5 to16-year-olds.

This methodology report details the procedures and protocols followed to ensure the HLS produces high quality, robust data. Specific analyses such as short fact sheets and reports can be accessed at http://www.hpa.org.nz/research-library/research-publications.

1.1 BACKGROUND

Prior to the introduction of the HLS in 2008, the Health Sponsorship Council (HSC) undertook a number of different monitor surveys to benchmark and monitor changes in New Zealanders' knowledge, attitudes and behaviour in response to its social marketing and health promotion programmes and community-level activities in the health sector. These included:

- Smokefree/Auahi Kore Monitor, which had been running since the early 1990s and had been run annually since 2003
- 2006/07 Gaming and Betting Activities Survey, which provided benchmark measures for the minimising gambling harm programme
- New Zealand Children's Food and Drinks Survey, undertaken in 2007 to provide benchmark measures for the nutrition and physical activity programme
- Sun Protection Triennial Survey, which monitored responses to the sun safety programme and had been undertaken since 1994

These monitors focused on adults, although the Gaming and Betting Activities Survey, the Children's Food and Drink Survey and the Sun Protection Triennial Survey also interviewed young people in the target age group for that particular programme.

In 2007, HSC reviewed the adult surveys and combined the majority of these into a single survey - the HLS.

---

[1] HPA is a New Zealand Crown entity formed in 2012 by the merger of the Health Sponsorship Council (HSC) and the Alcohol Advisory Council (ALAC), and some health promotion programmes previously delivered by the Ministry of Health.

1.2   OBJECTIVES OF THE HLS

The objectives of the HLS are to:

- measure progress against HPA's existing programme plans
- provide quality measures for Statement of Intent reporting requirements
- monitor short, medium and long-term societal changes in attitudes, knowledge and behaviours, and track changes in views about the social desirability and acceptability of various measures of tobacco control, minimising gambling harm, nutrition and physical activity, alcohol, sun safety, immunisation and mental health.

1.3   ETHICS

The 2014 HLS was voluntary and this was clearly explained to potential participants in HPA's brochure, on HPA's website, as well as verbally by the interviewer. The 2014 HLS was approved by the New Zealand Ethics Committee.

Confidentiality of all information provided by respondents in the interviews was assured by the Privacy Act 1993. The final, stored electronic records contain no identification of the participating respondents and responses can only be analysed as overall or grouped data.

2.   POPULATION AND FRAME

This section discusses the target population, the survey population and the sample frame.

The *target population* is the population the survey aims to represent. All statistics for the survey refer to the target population. The *survey population* is the population that had a probability of being selected to participate in the survey. For reasons discussed below, a small proportion of people did not have a chance of being selected to participate in the survey. As a result, the survey population is slightly smaller than the target population. The sample *weights* are designed to reflect the target population, so that the weighted statistics produced from the HLS can be taken to be representative of the target population.

The *sample frame* is the list of areas, and the lists of dwellings and people within these areas, that were used to select the HLS sample from the survey population.

2.1   TARGET POPULATION

The target population was the usually resident civilian population aged 15 years and over living in permanent private dwellings in New Zealand. The size of the target population was 3,263,184 individuals (Statistics New Zealand, 2014). This is the 2013 resident population adjusted to exclude those who do not reside in permanent private dwellings (based on information from the 2013 Census).

For reasons of practicality and cost-effectiveness, the target population is defined to include only permanent private dwellings, so temporary private dwellings are excluded, including caravans, cabins and tents in a motor camp, and boats. The target population also excludes non-private dwellings. Examples of non-private dwellings are hotels, motels, guest houses, boarding houses, homes for the elderly, hostels, motor camps, hospitals, barracks, and prisons.

People were eligible to be interviewed at their usual residence only. If they were temporarily visiting a household that was selected into the HLS they were not eligible for selection as part of that household. This process ensured that double counting was not possible.

People who were usually resident in a private dwelling in New Zealand, but who were temporarily overseas for some of the survey period, were included in the target population. In the majority of cases these individuals had a chance of being selected in the survey, as the survey provider made six call-backs to non-contacted households in the sample over the survey period. The benchmarks used in weighting the survey also included usual residents temporarily overseas.

2.2   SURVEY POPULATION

Households were not included if they were in meshblocks with fewer than nine occupied dwellings (according to the 2013 New Zealand Census of Population and Dwellings), or located off the main

islands of New Zealand (North, South and Waiheke), such as those on other sparsely inhabited off-shore islands, on-shore islands, waterways, and inlets. This meant that a small number of households (less than 2%) that were part of the defined target population were excluded from the survey population. However, these have been accounted for in the final estimates via the survey weights. Due to the small number of households omitted, any possible bias is likely to have little consequence.

2.3   SAMPLE FRAME

Meshblocks are the smallest geographical measure used by Statistics New Zealand. They vary in size from a city block to a large rural area and are used to make up other geographical measures in New Zealand (Statistics New Zealand, n.d.).

New Zealand 2013 Census meshblocks were used as part of an area-based frame of 37,525 meshblocks that had 10 or more dwellings. A sample of 350 meshblocks was selected from this frame, and these were the primary sampling units (PSU) of the HLS. Interviewers listed all the addresses in each of these areas. These lists of dwellings were then used as a frame from which a sample of dwellings was selected from each meshblock. One eligible adult and/or one parent/caregiver (if any) was then selected from each selected dwelling.

3.   DESIGN EFFECT

The net effect of a complex design can be measured by the design effect (or DEFF). The DEFF is the ratio of the variance (a measure of precision) of an estimate achieved by a complex design relative to the variance of the same estimate that would be achieved by a simple random sample of the same size. The closer the DEFF is to 1, the closer the design is to simple random sampling. Design effects of between 2 and 4 are typical in population health surveys, which means the variance is larger than would have been obtained using a simple random sample. A complex design like that used in the 2014 HLS is less precise than a simple random sample with the same sample size, but is much more precise than could be achieved by a simple random sample with the same budget.

Nevertheless, DEFFs should not be too large. On the one hand, it is appropriate for weights to vary across the sample, otherwise it would not be possible for Māori and Pacific peoples to have an increased chance of selection in the sample. On the other hand, if the variation in weights is too extreme, the DEFF will be very large, and this would be counter-productive for all statistics, even for Māori and other sub-population groups. The methods to sample sub-populations for the 2014 HLS were used to ensure the sample design was appropriate for achieving adequate precision for national and sub-population estimates within the survey budget.

Note that the design effects are different for each statistic. Table 3-1 (page 12) presents the design effects for a key indicator from each programme area. These are calculated by dividing the variance from the sample proportion by an estimate of the variance of an unrestricted sample with unknown parameters, as estimated from the HLS sample: $\frac{proportion \times (1 - proportion)}{sample\ size}$

Table 3-1: Design effects for four key indicators from the 2014 HLS for each sample, by ethnic group

Indicator	Ethnic group	General sample	Parent/Caregiver sample
Current smoker	Māori	1.85	2.14
	Pacific	2.87	1.74
	Asian	1.14	1.17
	European/Other	2.15	2.84
	Total	2.16	3.52
Sunburnt last summer	Māori	2.48	1.61
	Pacific	4.37	1.79
	Asian	1.28	9.72
	European/Other	1.79	2.94
	Total	2.15	6.14
Eats fruit at least twice a day	Māori	2.97	1.51
	Pacific	4.31	2.25
	Asian	2.26	5.14
	European/Other	2.09	3.47
	Total	2.79	6.76
Gambler	Māori	2.62	-
	Pacific	2.25	-
	Asian	3.02	-
	European/Other	3.37	-
	Total	3.51	-

Note: the parent/caregiver sample was not asked any questions from the gambling section of the questionnaire.

4. SAMPLE DESIGN

The HLS was designed to be able to produce nationally representative estimates. The 2014 HLS adopted a multi-stage, stratified, probability-proportional-to-size (PPS) of the meshblocks sampling design.

4.1 RATIONALE FOR THE SAMPLE DESIGN

A primary consideration in the sample design of the HLS was the need for sufficient samples of Māori, Pacific peoples, and people of European/Other ethnicities, as well as low socio-economic status groups and current smokers. The main group of interest was adults aged 15 years and over, but it was also important there be enough parents and caregivers of 5 to 16-year-olds to be able to analyse the results of this group with confidence.

The challenge for the sampling methodology was to arrive at a sample that could:

- provide national, projectable figures
- use a survey method with higher (face-to-face), rather than lower (phone, mail, web) public participation
- deliver 2,500 interviews with adults aged 15 years and over, made up of 450 interviews with Māori, 300 with Pacific peoples, and 250 with Asian people
- deliver 800 interviews from parents/caregivers of 5 to 16-year-olds (including interviews with 200 Māori, 200 Pacific peoples and 100 Asian people)
- provide the minimum design effect for the overall sample and specific target groups within the budget for the survey.

The simplest possible sample design would be a random sample drawn from all people in New Zealand, so that everyone has an equal and independent chance of being selected in the sample. However, a design of this type would not be feasible because:

- the sample would be geographically very dispersed, requiring interviewers to travel great distances between interviews
- it would not result in large enough numbers of Māori or Pacific peoples to enable adequate statistics for these groups.

Because of this, the 2014 HLS used a complex sample design.

Complex designs have two main features that affect the precision of statistics coming from the survey.

1. *Different people have a different chance of selection.* This was captured in the 'weight', which is the number of people that each survey respondent represents in the target population. In the 2014 HLS, Māori and Pacific peoples had lower weights than other people to reflect the fact that these groups had an increased chance of selection in the sample relative to simple random sampling. Sampling of one adult per household also led to different weights, because adults in larger households received a larger weight. In the 2014 HLS, the selection weight for adult participants who were selected for the

parent/caregiver sample was adjusted to account for their increased chance of selection in the adult sample.

2. *The sample was 'clustered'*. In the HLS a sample of meshblocks was selected, and then a sample of households was selected from each meshblock. If the households in the sample were shown on a map of New Zealand they would appear clumped. Clustering made the survey more cost effective as interviewers did not have to travel between as many areas as they would if simple random sampling was used.

4.2   SAMPLE SELECTION PROCEDURE

A four-step selection process was used to achieve the sample.

Step 1: Put all meshblocks into strata
Using Statistics New Zealand meshblocks as the initial unit of sample, two strata were formed – a Pacific people's stratum consisting of meshblocks in which 20% or more of the population were of Pacific ethnicity (based on total ethnicity), and another stratum consisting of all of the ("other") remaining meshblocks.

Step 2: Select meshblocks within strata
Meshblocks vary considerably in size and were, therefore, selected by PPS design within each stratum. The size measure was the number of occupied dwellings in the meshblock according to the 2013 Census. This means that larger meshblocks had an increased chance of selection in the design. In total, 350 meshblocks were drawn randomly, with 56 selected from within the Pacific stratum and 294 selected from the Other stratum.

Step 3: Select households within meshblocks
Within each meshblock, some households (on average 10, with a maximum of 15) were selected to form the core sample, and some households were selected to form the screened or booster sample that oversampled Māori and Pacific peoples.

Households in the core sample were selected by a systematic procedure of beginning at a random dwelling pre-allocated in the meshblock and knocking on the door of every k^{th}[2] house.

Up to 22 of the dwellings in between the kth houses were then selected as the screened sample. In up to 14 of these 22 dwellings, both Māori and Pacific peoples were eligible to be sampled, in the remaining eight dwellings only Pacific peoples were eligible to be sampled.

There was no substitution of households or respondents if the selected household or respondent was not contactable or was unavailable.

Step 4: Select respondents within households
The procedure for selecting respondents in the 'core' and 'screened' households was essentially the same (Figure 4-1, page 16).

---

2 K is determined by the number of dwellings in the meshblock. For example, in a small meshblock K might be every fifth dwelling, while in a large meshblock it might be every 10th dwelling.

Within each household, all eligible adults who were aged 15 years and over and usually resided at that dwelling were identified. The ethnicities of eligible respondents were obtained by proxy from the person who answered the door using the Statistics New Zealand question that was used in the 2001 and 2006 Censuses. The interviewer asked if any children aged 5 to 16 years usually lived four or more days per week in the household. If so, the interviewer recorded whether any of the adults were parents or caregivers of any children aged 5 to 16 years.

If there were no 5 to16-year-old children living in the household, then the household was included as part of the adult sample (which occurred approximately 50% of the time), and one adult was randomly selected.

If there were 5 to 16-year-old children living in the household then one parent/caregiver was randomly selected. To reduce the number of dwellings in which two interviews were required, the probability of selection of parent/caregivers for the adult sample was doubled. In some households a single person was interviewed both as part of the parent/caregiver sample and as part of the adult sample, while in other households two people were interviewed, one for the parent/caregiver sample and another for the adult sample.

Overall, 2,594 people aged 15 years and over participated in the adult sample and 742 people participated in the parent/caregiver sample.
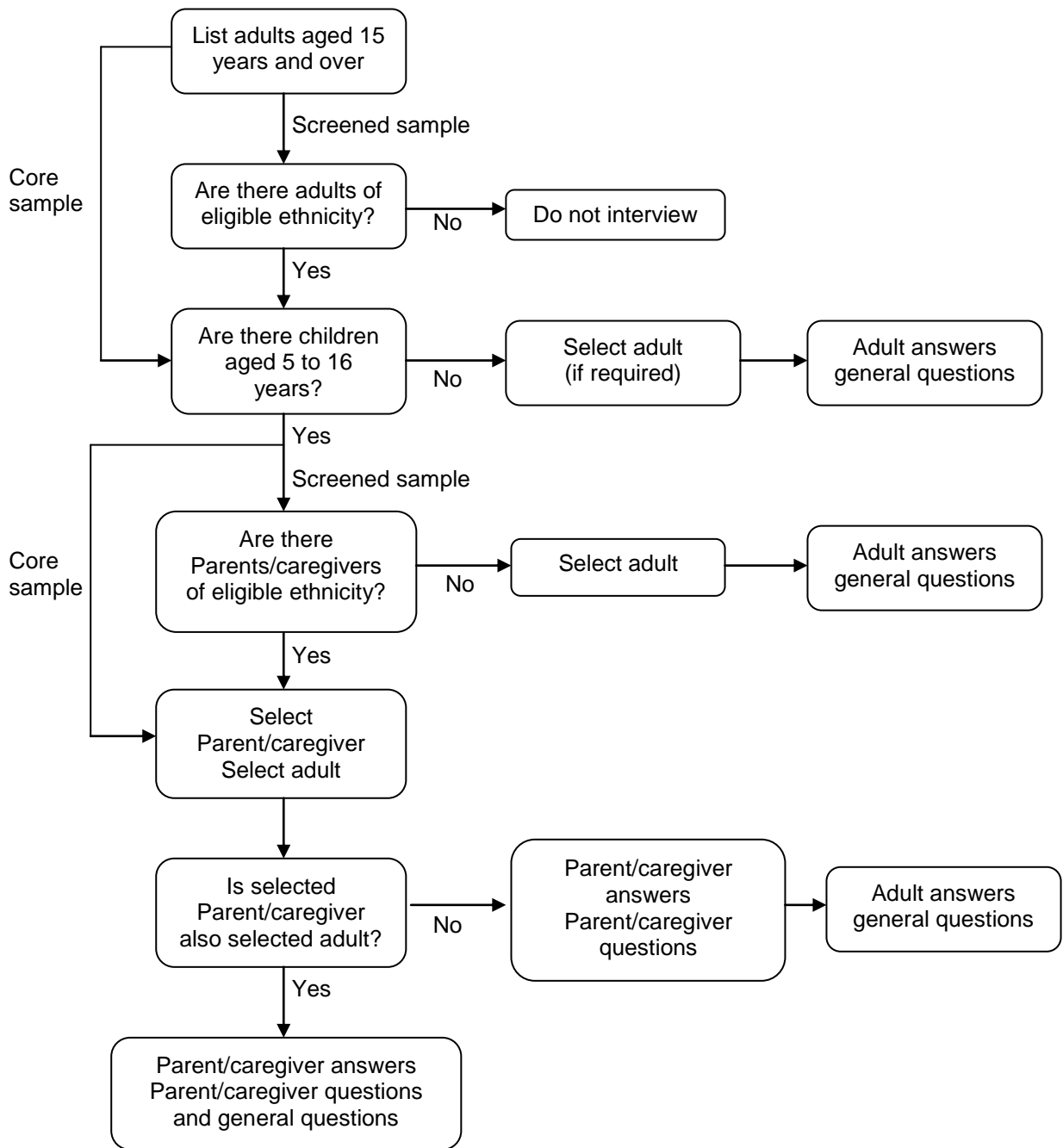
Figure 4-1: Diagram of the 2014 HLS respondent selection process within the household

5. DATA COLLECTION INSTRUMENTS

5.1 QUESTIONNAIRE CONTENT

The 2014 HLS questionnaire is available from http://www.hpa.org.nz/research-library/research-publications. Table 5-1 outlines the topic areas in the questionnaire.

The gambling section, with more than 60 questions, was the largest section of the questionnaire. The majority of these gambling questions were sourced from the 2006/07 Gaming and Betting Activities Survey and the Gambling Participation and Attitudes Survey to facilitate comparisons with data collected from these surveys and monitor any changes in gambling behaviour, knowledge and attitudes since HPA's minimising gambling harm programme was established.[3]

Other questions in the 2014 HLS were also sourced from previous surveys or pilot surveys. These included the 2008, 2010 and 2012 HLS, the Smokefree/Auahi Kore Monitor, the 2007 Children's Food and Drinks Survey, the 2010 Sun Exposure Survey, and the New Zealand Tobacco Use Survey.

The 2014 HLS questionnaire was informed by advice from HPA staff working in the specific programme areas, external researchers working in the specific topic areas, as well as other surveys.

Table 5-1: Summarised content of the 2014 HLS questionnaire

Programme area	Information domains	Output details
All	Demographics	- Age, gender, ethnicity (of adult, and child if applicable). - Immigrant status. - Employment status, leadership status, health sector status, workplace activity, highest qualification, household income. - Household composition.
	Re-contact	Respondents were asked if they would consent to be re-contacted to participate in further HPA research. Details from the re-contact question responses have been kept separately from the main dataset to maintain confidentiality.
Lifestyle	Sedentary behaviour	- Time spent watching television. - Internet and social media use. - Access to Internet.
	Transport	Mode of transport to main weekly activity (of adult, and child if applicable).

---

3 The HSC previously established the problem gambling programme.

Programme area	Information domains	Output details
	Neighbourhood	• Safety of local neighbourhood (for children).
Sun safety	Sun protection-related demographics	• Skin type (of adult, and child if applicable). • Workplace sun safety policy.
	Sun protection behaviour	• Use of sun protection behaviours (of adult, and child if applicable). • Tanning behaviour. • Skin checks.
	Incidence of sunburn	• Incidence of mild and extreme sunburn last summer (of adult, and child if applicable).
	Campaign monitoring	• Recognition and understanding of the Sun Protection Alert.
Healthy eating	Healthy eating-related demographics	• Consumption of different food types (by adult and child). • Main food provider status.
	Healthy eating behaviour	• Main meal preparation and child involvement. • Meal planning. • Agreement scale – changing household consumption of full sugar drinks, cost of full sugar drinks.
	Shopping patterns	• Weekly spend on food and drinks from supermarket-type locations, green grocer, fruit and vegetable shops or markets, farmers' markets, and from convenience-type locations.
	Attitudes	• Fastfood sponsorship. • Television fastfood advertising (for children). • School environment food consumption (for children).
	Campaign monitoring	• Monitor behaviour related to 'Breakfast-eaters have it better'.
Tobacco control	Tobacco control-related demographics	• Smoking status. • Intention to smoke in the future. • Stages of nicotine addiction. • Heavy smoking index. • Smoking around children in the home and cars (for children).
	Quitting	• Resources used. • Know where to seek help. • Nicotine replacement medications.
	Exposure	• Cigarette or tobacco packs displayed. • Brand recognition.

Programme area	Information domains	Output details
	Knowledge	• Knowledge of how many adult smokers there are in New Zealand. • Knowledge of government smoking rates reduction by 2025.
	Attitudes	• Attitudes towards smoking in a number of indoor and outdoor settings. • Smoking in New Zealand. • Attitudes towards regulation of smoking. • Attitude towards regulation of cigarette or tobacco sales. • Tobacco sales to minors. • Nicotine content of cigarettes.
	E-cigarettes	• Usage and attitudes towards use. • Helpfulness in assisting to quit smoking tobacco.
	Marijuana	• Usage.
Gambling harm	Gambling harm-related demographics	• Participation in gambling activity - nature and frequency of this participation. • Stages of gambling addiction. • Personal gambling harm (Problem Gambling Severity Index).
	Exposure	• Gambling advertising and perceived behaviour relating to this. • Gambling harm of a significant other. • Household gambling harm. • More time or money spent on gambling than wanted (self and other). • Gambling harm service use. • Strategies used to avoid gambling harm. • Self-monitoring of gambling behaviour. • Interaction with staff at gaming machine venues. • Harmful gambling information at gaming machine venues.
	Awareness	• Gambling harm advertising. • Signs of harmful gambling. • What to do to help someone with a gambling problem. • Early signs of harmful gambling. • Services available. • Legal requirements of gaming machine venues.
	Attitudes	• Social undesirability of gambling activities. • Gaming machines in bars/clubs. • Concern towards level of gambling in community.

Programme area	Information domains	Output details
Alcohol	Alcohol-related demographics	• Drinking status.
	Attitudes to regulation changes	• Hours. • Purchasing age. • Advertising, promotion and sponsorship. • Number of outlets for alcohol purchase in local area.
	Exposure	• Sources of alcohol advertising exposed to in the past three months.
	Alcohol-related injuries	• Accident in the past 12 months and alcohol consumption in previous 6 hours. • Perception of consumption of alcohol contributing to injury.
Physical activity	Physical activity behaviour	• Measure of physical activity level through type and frequency of physical activity. • Sedentary activity frequency. • Time spent playing/practising sport or other exercise activities outside of school hours (for children). • Frequency of activities in summer participated in by household (for children).
Other HPA areas	Immunisation	• Child vaccination history. • Concerns about recommended childhood vaccines. • Flu vaccination status and attitudes. • Eligibility for free flu vaccine.
	Mental health	• Depression screening. • Knowledge of depression. • Mental health stigma and discrimination. • Life stress and isolation.
	General health	• Connectedness. • Cultural identity. • Weight and height. • Perception of own weight. • Primary healthcare. • Internet use for health information.

6. DATA COLLECTION

6.1 COLLECTION MODE

Interviews were conducted in respondents' homes. Interviewers entered responses directly into laptop computers, with some questions being completed by the respondents independently. Show cards with predetermined response categories were used to assist respondents where appropriate.

6.2 ENUMERATION

Household from meshblocks were pre-selected for inclusion in the survey using the New Zealand Post address database. Each meshblock was re-enumerated when the interviewer first visited, in order to record new dwellings built and those removed since the last pre-Census enumeration and release of the New Zealand Post address list. The details of the new dwellings were entered into CBG's 'Sample Manager' software while the interviewer was in the field, allowing these households to be included in the random selection process for the meshblock.

6.3 CALL PATTERN

The 'call' refers to one visit on one day during a particular time period. Up to 10 calls to each sampled dwelling were made at different times of the day and on different days of the week, before accepting that a dwelling was a non-contact. Calls were recorded as unique events only if they were made at least two hours apart. Calls were spread out over the duration of the fieldwork. Six calls were made in the survey month in which the meshblock was issued. If no contact had been achieved by this point, there was a pause with no attempted contact with the dwelling for one to two weeks, before attempting four more calls. For 92% of households, the first (or only) interview took place within seven calls (Figure 6-1, page 22). 'Closed' meshblocks were revisited during a mop-up phase ie, when visiting households where no contact had been established or the selected respondent was unable to take part at that time but did not refuse to participate.
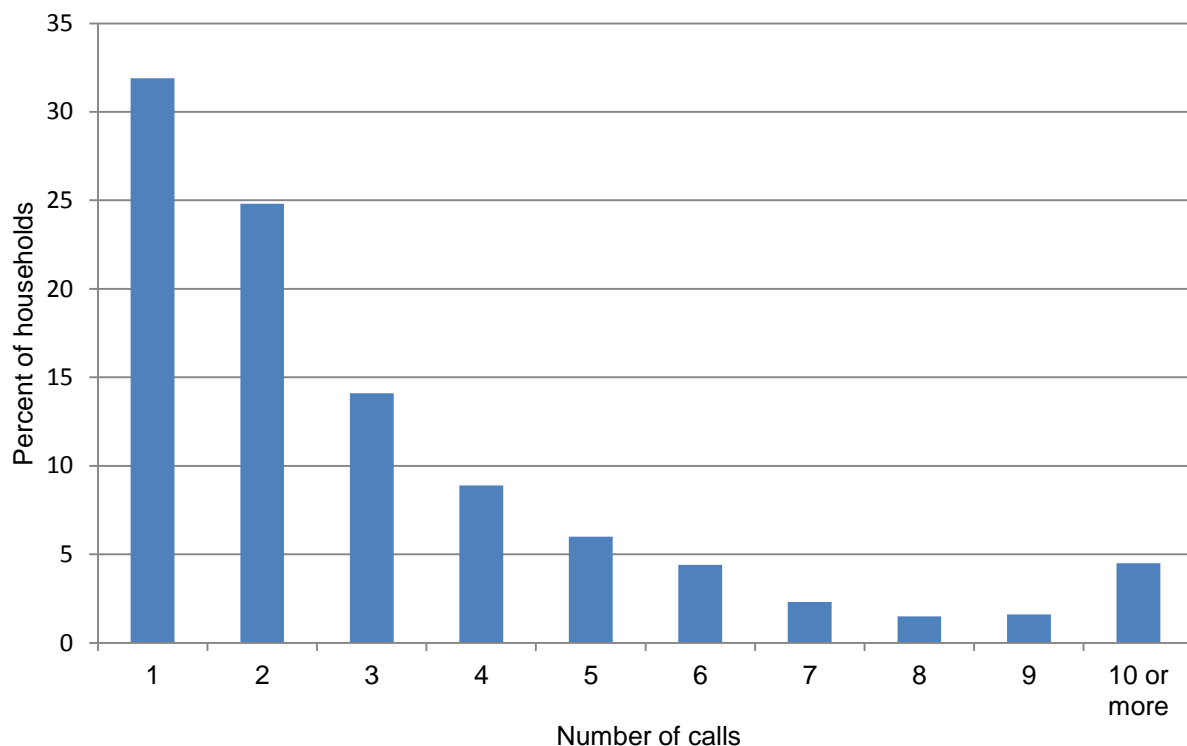
Figure 6-1: Proportion of households agreeing to first interview, by number of calls, 2014 HLS

6.4 PERFORMANCE AND QUALITY CONTROL

Interviewers were monitored by CBG management by:

- in-field assessment to ensure survey protocols were being followed correctly
- examination of individual performance metrics and exploration of strategies to improve these if necessary
- checking of a random selection of completed interviews by phoning respondents to confirm the interview was completed according to survey protocols and to collect satisfaction ratings.

Participants were also left with feedback postcards that they could use to send feedback directly to CBG, anonymously if they chose. In addition, CBG operated a Freephone survey helpline that participants could call if they had any questions about the survey or wanted to provide feedback.

The results of these quality checks were communicated to the individual interviewers on a regular basis throughout the fieldwork period, with additional training and mentoring provided where required.

6.5 INFORMED CONSENT

The 2014 HLS was voluntary. Consent was obtained without coercion. No incentive was offered.

Selected households were mailed an invitation letter and information brochure prior to the interviewer's first visit. Participants selected for the survey were presented with a copy of these documents as part of the informed consent process. Participants were asked to sign an electronic consent form and were given a copy of the consent form to keep. The consent form included a request for an interpreter if required (in a range of different languages, including New Zealand Sign Language), and the option was available to match respondents and interviewers by ethnicity and gender, although this was not requested. The information brochure, as well as the translations and further questions and answers were all available on the HPA website for respondents to view.

6.6 PILOT

A pilot survey of 100 respondents was completed from 26 March to 6 April 2014. The pilot was designed to mimic the main study in order to test:

- the duration of each survey type and the sections within
- that the questionnaire loaded into the CAPI software performed as expected and electronic sample management behaved as expected
- wording of new questions and how respondents understood them
- flow of the questionnaire
- that questions would provide useful information
- that interviewer training was appropriate and adequately prepared them for fieldwork
- that interviewer materials and resources were fit for purpose.

The survey design and sampling method had already been successfully used for the 2008, 2010 and 2012 HLS.

The pilot sample was not random, as people were selected to represent the different mix of ethnic groups, age groups, and geographic locations likely to be included in the main survey (a purposive sample). Once the pilot was reviewed, a number of questions were removed from the questionnaire, or further refined.

6.7 FIELD DATES

Interviews for the main survey were conducted from 5 May to 10 August 2014.

6.8 RESPONDENT BURDEN

HPA sought to minimise the burden on respondents by:

- seeking interviews by appointment rather than requesting immediate participation
- reducing the number of dwellings where two interviews were required, by increasing the probability of the randomly selected parent/caregiver also being the randomly selected adult

- planning for a 50-minute average duration. In practice, a duration of 54 minutes[4] eventuated for adults and 57 minutes for parent/caregivers. Where the parent/caregiver was also the selected adult, therefore answering both sets of questions, the average duration was 73 minutes. Two interviews were conducted in 230 dwellings, one with a parent/caregiver and one with another adult. In these dwellings, the combined average interview duration was 90 minutes

- using showcards wherever possible to assist answering

- inviting open-ended answers to enable respondents to feel they could express themselves, rather than being simply an information source.

---

[4] These times are the CAPI times and include all question modules. They do not include the time spent in a household before or after the interview was conducted.

7.   RESPONSE RATES

The main measure used to assess the overall quality of a survey is the response rate. The response rate is a measure of how many of the people who were selected to take part in the survey actually participated. The response rate reflects the proportion of people interviewed from those who were selected into the sample, and describes the success of the study in terms of achieving cooperation from the population being measured. A high response rate means the survey results are more representative of the target population.

There are four components to the response rate calculation:

- ineligibles eg, vacant sections, vacant dwellings, non-residential dwellings and those not available during the survey period

- respondents (interview conducted, respondent confirmed to be eligible for the survey)

- eligible non-respondents (interview not conducted, but enough information collected to indicate that the household did contain an eligible adult)

- unknown eligibility[5] eg, non-contacts and refusals who provided insufficient information to determine eligibility ie, households in the screened samples.

The 2014 HLS response rate was calculated as follows:

$$\text{Response rate} = \frac{\text{number of respondents}}{\begin{bmatrix} \text{number of} \\ \text{respondents} \end{bmatrix} + \begin{bmatrix} \text{number of eligible} \\ \text{non- respondents} \end{bmatrix} + \begin{bmatrix} \text{estimated number of eligibles} \\ \text{from the unknowns} \end{bmatrix}} \times 100$$

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 contact could not be made with the estimated number who would be eligible, they were classified as non-respondents.

The estimated number of unknown eligibles was calculated as follows:

$$\begin{bmatrix} \text{Estimated number of eligibles} \\ \text{from the unknowns} \end{bmatrix} = \begin{bmatrix} \text{number of} \\ \text{unknowns} \end{bmatrix} \times \frac{\begin{bmatrix} \text{number of} \\ \text{respondents} \end{bmatrix} + \begin{bmatrix} \text{number of eligible} \\ \text{non- respondents} \end{bmatrix}}{\begin{bmatrix} \text{number of} \\ \text{respondents} \end{bmatrix} + \begin{bmatrix} \text{number of eligible} \\ \text{non- respondents} \end{bmatrix} + \begin{bmatrix} \text{number of} \\ \text{ineligibles} \end{bmatrix}}$$

For the adult and the parent/caregiver samples a separate response rate was calculated for each primary sampling unit (PSU). This was then adjusted to the estimated number of eligible households in that PSU. Once this was done the average response rate across all of the PSUs was calculated.

Unweighted response rates are calculated using the raw counts and reflect the success of the survey in terms of being able to get the people selected to participate.

---

5 This grouping applies to the response rate calculated for parent/caregivers. The response rate calculated for adults has all these outcomes added to the eligible non-respondents category.

Weighted response rates take probability of selection into account and reflect the success of the survey in terms of the population being measured. These have been used for the HLS because of the sample design and reflect that different dwellings had a different chance of selection due to screened samples being used to boost the proportions of Māori and Pacific peoples in the survey.

7.1 ADULT SAMPLE RESPONSE RATE

The unweighted response rate for the adult sample was 76.4%, compared with 83.1% for the 2012 HLS, 56.7% for the 2010 HLS and 63.7% for the 2008 HLS.

7.2 PARENT/CAREGIVER SAMPLE RESPONSE RATE

The unweighted response rate for the parent/caregiver sample was 82.7%, compared with 87.7% for the 2012 HLS, 54.8% for the 2010 HLS and 63.2% for the 2008 HLS.

8. DATA PROCESSING

This section outlines the processes used to collect, check and output the data for the 2014 HLS.

8.1 DATA CAPTURE

Questionnaire responses were entered directly onto interviewers' laptops. As interviewing progressed, completed interviews were uploaded to CBG's data server, from where they were compiled for inspection, coding and editing. Interviews were uploaded to the server by each interviewer on every day they were active in the field.

8.2 CODING

Different types of questions were used in the 2014 HLS. Single-response closed-ended questions, which a respondent can only give one response to, were coded as is. Some questions allowed for multiple responses. For these questions all responses were retained, with each response shown as a separate variable on the data file.

Open-ended questions were used extensively. For these, the interviewer keyed in the verbal answers, as near as possible to the respondent's spoken words. Coding of these was then done by HPA's data processing team.

Coding of open-ended questions was undertaken by initially collating answers given by respondents to each open-ended question. These answers were examined jointly by the researcher and a data specialist to search for recurring points or themes. Each recurring point/theme was identified as a code. All answers falling sufficiently close to that point/theme ie, differing only in the words the person used to describe it, were assigned to that code. Note that where an open-ended question was sourced from a prior HPA survey, the code frame used previously was also used for the 2014 HLS when appropriate, to enable comparisons between the surveys.

Questions with an "Other, please specify" code were treated in the same way as open-ended questions. In this case, the number of original codes was extended to accommodate any further recurring answers. In some instances, interviewers tend to put into "Other, please specify" an answer that fits into one of the pre-coded categories. In this case, the answer was assigned that code.

All open-ended responses have been retained, to inform any further review of the codeframes used.

8.3 SECURITY OF INFORMATION

Any information collected in the survey that could be used to identify individuals has been treated as strictly confidential. Data were transferred from interviewers' laptops to head office at CBG by a secure internet upload facility.

Names and addresses of people and households who participated in the survey have been stored separately from the response data.

8.4 IMPUTATION

A small number of respondents (less than 0.5%) did not answer their age. However, all of these provided an age group so age was imputed as the midpoint of this range where needed for specific age analysis. For those who selected the 65+ age group, age was imputed by randomly selecting another respondent with the same gender, ethnic group, employment and education status.

Income was missing for 17.3% of respondents in the General sample. However, 1.1% were able to be imputed using parent/caregiver responses from the same household. This was also done for a very small number of missing responses for the food and drink expenditure questions.

8.5 CREATION OF DERIVED VARIABLES

A number of derived variables have been created for the 2014 HLS dataset.

Ethnicity

Ethnicity was calculated using prioritisation, where each person is allocated to a single ethnic group based on the ethnicities they have identified with, in the prioritised order of Māori, Pacific peoples, Asian, and European/Other (Ministry of Health, 2004). For example, if someone identifies as being Chinese and Māori, under the prioritised ethnic group method, they are classified as Māori for the purpose of analysis. The way that the ethnicity data is prioritised means that the group of prioritised European/Other effectively refers to non-Māori, non-Pacific, and non-Asian people. Prioritisation is a method outlined in the Ethnicity Data Protocols for the Health and Disability Sector as a useful method for grouping people into independent ethnic groups for analysis (Ministry of Health, 2004).

Note that as ethnicity was collected as a multiple response variable it is possible to also analyse it using total response or sole/combination methods.

Smoking status

The definitions used for smoking status areas follows:

- *Never smoker:* has never smoked tobacco.
- *Past experimental:* has ever smoked tobacco, but never started smoking [regularly].

- *Current smoker:* has ever smoked tobacco, and now smokes at least once a month or more often.

- *Recent/past quitter:* has ever smoked tobacco, but has now stopped smoking.

Gambling type

Gambling types are often classified into two categories, those where winnings can be immediately 'reinvested' and those where they cannot. The former referred to as 'continuous' and the latter 'non-continuous' (Abbott & Volberg, 1996). For the HLS these two groupings were combined with frequency in the same way they were presented for the 2006/07 Gaming and Betting Activities Survey (National Research Bureau, 2007).

- *Non gamblers:* did not participate in any gambling activities in the last month.

- *Infrequent gamblers:* participated in any gambling activities less than once a week.

- *Frequent, non*-continuous gamblers: participated weekly or more often in non-continuous[6] forms of gambling.

- *Frequent, continuous gambler:* participated weekly or more often in continuous[7] forms of gambling.

Neighbourhood socioeconomic deprivation

The New Zealand Index of Socioeconomic Deprivation 2013 (NZDep2013) has been linked to the 2014 HLS as a measure of neighbourhood socioeconomic deprivation and a proxy for individual socioeconomic position. The NZDep2013 was created using nine variables[8] from the 2013 Census data, with a decile value calculated for each meshblock (Atkinson, Salmond & Crampton, 2014). For some analyses of the 2014 HLS, these deciles have been grouped, so that deciles 1–3 are referred to as low deprivation, 4-7 as moderate (or mid) deprivation, and 8-10 as high deprivation.

Household equivalised income

To measure household income, respondents were asked to choose an income range that represented their total household income from all sources before tax in the previous 12 months. However, household income by itself is not always an accurate measure of living standards as, for example, a two-person household with a total household income of $100,000 is likely to be quite different in many characteristics from that of a six-person household with a total household income of $100,000. Therefore, equivalised household income was derived using the revised Jensen Index (Jensen, 1988). The revised Jensen Index is a recognised equivalisation index used within New Zealand (Blakely, 2002; Ministry of Health, 2010), that takes into account the number of adults, the

---

6 Non-continuous forms of gambling include lottery games, going to casino evenings/buying raffle tickets for fundraising, participating in sweepstakes, making bets with family/friends and other gambling activities.

7 Continuous forms of gambling include playing electronic gaming (pokie) machines, betting on horse or dog races, or sports events, table games at casinos, housie and bingo, mobile phone games for money, online activities for money or prizes through an overseas website.

8 Receiving a means-tested benefit, low household income, not owning the home you live in, single-parent family, unemployment, no school qualifications, household overcrowding, no access to internet at home and no access to a car.

number of children (younger than 18-years-old) and the ages of the children living in the household.

Income was calculated as the mid-point of the band the respondent selected. If the respondent did not provide a band, but another person in the household was also interviewed and did provide a band ie, different adults were interviewed for the parent/caregiver and the adult sample, then the band selected by the other person in the household was used. If the respondent selected the band 'Over $250,000', then $275,000 was used as the household income. Some respondents did not give an answer using the narrower bands first provided to them, so were asked the question again using wider income bands. If these respondents selected the wider band of '$100,000-$250,000', their income was calculated as $150,000 based on the mean of the mid-points of the three narrower bands between $100,000 and $250,000.

Household income was divided by the formula developed by Jensen:

$$Income\ equivalence\ of\ a\ household = \frac{[(\text{number of adults aged 18+}) + (w_1 \times \text{number of children}) + (w_2 \times \text{the sum of the ages of all the children})]^u}{2^u}$$

Where w1 = 0.460697, w2 = 0.0283848 and u = 0.621488. The mid-points of the ranges provided for the children's ages were used in this equation.

Equivalised household income was then divided into tertiles ie, three equal groups, of low, medium and high for use in some analyses.

Household equivalised expenditure on food and drinks

Respondents were asked how much money their household usually spends each week on food and drinks from different vendors. These variables have the same limitation mentioned above for household income, and Jensen's formula can also be used for expenditure (Jensen, 1988). The same process was followed to calculate household equivalised expenditure on food and drinks as was used to calculate household equivalised income (please see the description of this outlined above), with $425 used as the midpoint for the band '$401 or more'.

9. WEIGHTING

9.1 OVERVIEW OF WEIGHTING PROCESS

Most national surveys have complex survey designs, where different groups have different probabilities of being selected in the survey. These complex designs are used for a variety of purposes, including:

- reducing interviewer travel costs by ensuring the sample is geographically clustered or 'clumped'
- ensuring all sub-populations (especially the Māori and Pacific populations) have a sufficient sample to enable adequate estimates.

To ensure no group is under- or over-represented in estimates from a survey, a method of calculating estimates that reflects the sample design must be used. Estimation weights are used to achieve this, and can be thought of as the number of people in the population represented by a given survey participant. A weight is calculated for every respondent and these weights are used to calculate 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 eg, Māori and Pacific populations who are included in the booster samples receive lower weights. Also, groups that have a lower response rate eg, older men are usually assigned a higher weight so that these groups are correctly represented in all estimates from the survey.

Weights are designed 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.

9.2 PROBABILITY OF SELECTION WEIGHTS

The probability of selection for each respondent comes from three factors:

1. *The probability of the meshblock being selected.*

    For the 2014 HLS this was:

$$no.\,of\,meshblocks\,in\,the\,stratum \times \frac{no.\,of\,dwellings\,in\,the\,meshblock\,recorded\,in\,the\,2006\,Census}{Total\,no.\,of\,dwellings\,in\,the\,sampled\,meshblocks\,for\,the\,stratum}$$

For the Pacific stratum the number of meshblocks was 56 and the total number of dwellings in the stratum was 111,429.

For the Other stratum the number of meshblocks was 294 and the total number of dwellings in the stratum was 1,417,863.

2. *The probability of their dwelling being selected within the meshblock.*

For the 2014 HLS this was:

$$\frac{No.\ of\ dwellings\ with\ eligible\ respondents}{No.\ of\ private\ dwellings\ in\ the\ meshblock\ at\ the\ time\ the\ meshblock\ was\ sampled}$$

3. *The probability of the respondent being selected from all the eligible individuals within the dwelling.*

For the parent/caregiver sample this was:

$$\frac{1}{No.\ of\ parents\ in\ the\ household}$$

For the adult sample this was:

$$\frac{1}{2\ \times No.\ of\ adults\ in\ the\ household}$$

For the non-parents in the core sample this was:

$$\frac{1}{No.\ of\ adults\ in\ the\ household}$$

For the non-parents in the screened sample this was:

$$\frac{2}{No.\ of\ adults\ in\ the\ household + 1}$$

For parents/caregivers interviewed for both the parent/caregiver and the adult sample this was:

$$\frac{1}{No.\ of\ adults\ in\ the\ household + 1}$$

For adults (parent/caregivers or non-parents) interviewed for the adult sample (parent/caregiver interview done with someone else) the average probability of selection is the product of these three probabilities.

The selection weight applied to each respondent in the dataset is the inverse of the probability of selection for that respondent.

9.3 NON-RESPONSE ADJUSTMENT

Each selection weight was adjusted using the response rate of the meshblock the respondent was selected from. This adjustment was done to compensate for any non-response bias that may have arisen from people refusing to participate in the survey. The adjustment was made by dividing the selection weight by the response rate. Applying this adjustment at the meshblock level accounted for any bias that may have arisen due to differences at the area level eg, differing levels of deprivation in different meshblocks.

9.4 BENCHMARK POPULATIONS USED FOR THE 2014 ADULT SAMPLE

Benchmarking is an adjustment that ensures the proportion of particular groups in the sample match the proportions observed in the actual population estimates based on the Census data. The benchmarks used in the 2014 HLS weighting of the adult sample were population counts by:

- age group (15 to 24 years, 25 to 34 years, 35 to 44 years, 45 to 54 years, 55 years and over)
- gender (male, female)
- ethnic group (Māori, Pacific, Asian, European/Other)

Age, gender and ethnicity were included because these variables are related to health behaviour and to non-response and were a key output classification for the survey.

The most recent New Zealand Census was conducted in March 2013. The population benchmarks were calculated using the 2013 Census counts for usual residents.

The ethnic group counts from the Census were calculated using prioritised ethnic groups in order of Māori, Pacific peoples, Asian and European/Other (refer section 8.5).

Adjusting the selection weight with the benchmark weight helps remove any differences between the proportions of different groups in the sample compared to these proportions in the New Zealand population.

The 2014 HLS weights were adjusted back down to the sample size of the survey.

9.5 REPLICATE WEIGHTS

Standard errors are a measure of the precision of an estimate and replicate weights are a method for obtaining standard errors for any weighted estimate. In the 2014 HLS, jackknife replicate weights were used as part of the survey estimation procedures in the Stata version 13 statistical software package.

To remove bias in the estimate from any particular PSU 'delete-a-group' jackknife is used. This means that the estimate is first calculated from a sample of all respondents except those in a PSU,

and then this calculation is repeated excluding a different PSU each time. The standard error of the population estimate is based on the variation of the replicate estimates. For technical information on replicate variance estimation in surveys, see Rao and Wu (1988) and Shao and Tu (1995).

9.6   SURVEY ESTIMATES

Proportions

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

Proportions within population groups

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

Totals (counts)

Estimates of totals (counts) are given by the sum of the respondents of the weight multiplied by the variable of interest. For example, the estimate of the total number of people who smoke daily in the whole population would be given by the sum, over all respondents, of the number of respondents who smoke daily multiplied by the weight.

Averages (means)

The population averages eg, the average estimate of the number of smokers in New Zealand, are estimated by calculating the sum, over all respondents, of the weight multiplied by the variable of interest divided by the sum of the weights.

Averages within population groups

Sometimes the average within a group is of interest eg, the average estimate of the number of smokers in New Zealand among males. The estimate is given by calculating the sum, over respondents, in the group of the weight multiplied by the variable of interest, divided by the sum of the weights of respondents in the group.

10. TECHNICAL NOTES FOR ANALYSIS

The descriptive 2014 HLS analyses are presented in a series of fact sheets called *In Fact*. These use a number of specific techniques that are discussed below.

10.1 SUPPRESSION DUE TO SMALL NUMBERS

To ensure the survey data presented are reliable and that the confidentiality of the participants is protected, data are only presented when there are at least 30 respondents in the denominator (the population group being analysed). This ensures that no participant can be identified from the results.

10.2 CONFIDENCE INTERVALS

Ninety-five percent confidence intervals have been used to represent the sample error for estimates. A 95% confidence interval means there is a 95% chance the true value of the estimate (if the whole population was sampled) 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. However, even when there are overlapping confidence intervals the difference between the groups can be statistically significant, when the variance is sufficiently small.

Any differences between two variables where the confidence intervals overlapped were tested using the most appropriate statistical test for that data. The significance of many different statistical tests is represented by a probability value, or p-value. If a p-value is below 0.05, then we are 95% confident the difference between the two estimates is not due to chance.

11. DISSEMINATION OF DATA

There are several ways to access the results and data from the 2014 HLS:

- publications
- confidential microdata.

11.1 PUBLICATIONS

In Fact are information sheets highlighting interesting points from specific research. *In Fact* is designed to meet the needs of researchers, academics and people working in the health sector.

In Fact reports using data from the 2008, 2010, 2012 and 2014 HLS are available on the HPA website at: http://www.hpa.org.nz/research-library/research-publications.

Further publications using 2014 HLS data are planned and will be available from the same location.

11.2 ACCESS TO CONFIDENTIAL MICRODATA

The analyses presented in HPA publications are only a small proportion of those that could be undertaken. Confidentialised microdata from the 2014 HLS may be available by late 2015 for approved researchers to use for specific research projects.

The microdata will have all identifying information about individuals removed and be modified to protect individual information. Approval will be subject to certain criteria, terms and conditions and the researcher's organisation will have to sign an access agreement with HPA.

Contact HPA for more information
email: research@hpa.org.nz
phone: 64 4 917 0060

REFERENCES

Abbot, M. W. & Volberg, R. A. (1996). The New Zealand national survey of problem and pathological gambling. *Journal of Gambling Studies, 12*(2), 143-160.

Atkinson, J., Salmond, C. & Crampton, P. (2014). *NZDep2013 index of deprivation.* Wellington: Department of Public Health, University of Otago.

Blakely, T. (2002). *The New Zealand Census-Mortality Study: Socioeconomic inequalities and adult mortality 1991-94.* Wellington: Ministry of Health.

Jensen, J. (1988). *Income equivalences and the estimation of family expenditures on children.* Wellington: Department of Social Welfare (unpublished).

Ministry of Health. (2010). *Living standards and health: New Zealand 2006/07.* Wellington: Ministry of Health.

Ministry of Health. (2004). *Ethnicity data protocols for the health and disability sector.* Wellington: Ministry of Health.

National Research Bureau. (2007). *2006/07 Gaming and Betting Activities Survey: New Zealanders' knowledge, views and experiences of gambling and gambling-related harm.* Auckland: National Research Bureau.

Rao, J. N K. & Wu, C. F. J. (1988). Resampling inferences with complex survey data. *Journal of the American Statistical Association, 83*(401), 231-241.

Shao, J. & Tu, D. (1995). *The jackknife and bootstrap.* New York: Springer-Verlag.

Statistics New Zealand. (2014). *Dataset: Occupied dwelling type by age group of usual residents, for usual residents in occupied dwellings, 2013 Census.* Retrieved from http://nzdotstat.stats.govt.nz/wbos/index.aspx

Statistics New Zealand. (n.d.). *Geographic definitions.* Retrieved from http://www.stats.govt.nz/Census/2013-census/info-about-2013-census-data/2013-census-definitions-forms/definitions/geographic.aspx

APPENDIX A   SAMPLE SIZES

Tables A-1 to A-8 show the 2014 HLS actual sample sizes and the weighted counts by gender age, ethnicity, and NZDep2013 quintile for the adult and parent/caregiver samples.

Table A-1:   Sample sizes, by gender, 2014 HLS adult sample

Gender	Actual sample size	Weighted sample size
Males	1,086	1,242.3
Females	1,508	1,351.7
Total	2,594	2,594

Table A-2:   Sample sizes, by gender, 2014 HLS parent/caregiver sample

Gender	Actual sample size	Weighted sample size
Males	253	-
Females	489	-
Total	742	-

Table A-3:   Sample sizes, by ethnic group and gender, 2014 HLS adult sample

Ethnic group (prioritised response)	Gender	Actual sample size	Weighted sample size
Māori	Males	224	150.5
	Females	340	172.0
Pacific	Males	157	66.4
	Females	236	71.9
Asian	Males	96	140.6
	Females	121	156.7
European/Other	Males	609	884.8
	Females	811	951.5

Table A-4: Samples sizes, by ethnic group and gender, 2014 HLS parent/caregiver sample

Ethnic group (prioritised response)	Gender	Actual sample size	Weighted sample size
Māori	Males	69	-
	Females	149	-
Pacific	Males	58	-
	Females	113	-
Asian	Males	28	-
	Females	48	-
European/Other	Males	98	-
	Females	179	-

Table A-5: Sample sizes, by age group and gender, 2014 HLS adult sample

Age group	Gender	Actual sample size	Weighted sample size
15-24 years	Males	147	225.4
	Females	146	220.7
25-34 years	Males	158	186.8
	Females	276	204.4
35-44 years	Males	179	207.0
	Females	299	234.0
45-54 years	Males	195	222.0
	Females	256	241.8
55-64 years	Males	152	156.4
	Females	194	169.1
65+ years	Males	255	244.9
	Females	337	281.7

Table A-6:   Samples sizes, by age group and gender, 2014 HLS parent/caregiver sample

Age group	Gender	Actual sample size	Weighted sample size
15-24 years	Males	6	-
	Females	14	-
25-34 years	Males	52	-
	Females	151	-
35-44 years	Males	111	-
	Females	196	-
45-54 years	Males	68	-
	Females	107	-
55-64 years	Males	13	-
	Females	16	-
65+ years	Males	3	-
	Females	5	-

Table A-7:   Sample sizes, by NZDep2013 group and gender, 2014 HLS adult sample

NZDep2013 group	Gender	Actual sample size	Weighted sample size
Low (least deprived neighbourhoods)	Males	248	393.0
	Females	283	437.1
Mid	Males	422	560.3
	Females	568	574.1
High (most deprived neighbourhoods)	Males	411	288.7
	Females	639	340.6

Table A-8:   Sample sizes, by NZDep2013 group and gender, 2014 HLS parent/caregiver sample

NZDep2013 group	Gender	Actual sample size	Weighted sample size
Low (least deprived neighbourhoods)	Males	63	-
	Females	78	-
Mid	Males	89	-
	Females	157	-
High (most deprived neighbourhoods)	Males	101	-
	Females	254	-