Canadian Community Health Survey (CCHS)
Annual component

User guide
2012 and 2011-2012 Microdata files

June 2013

Statistics Canada Statistique Canada

Canada

TABLE OF CONTENTS

WHAT'S NEW IN THE 2012 CANADIAN COMMUNITY HEALTH SURVEY? 1

1. INTRODUCTION 4

2. BACKGROUND 5

3. CCHS REDESIGN IN 2007 6

4. CONTENT STRUCTURE OF THE CCHS 7

4.1 Common content 7
4.2 Optional content 8
4.3 Rapid response content 8
4.4 Content included in data files 8

5. SAMPLE DESIGN 11

5.1 Target population 11
5.2 Health regions 11
5.3 Sample size and allocation 11
5.4 Frames, household sampling strategies 13
5.4.1 Sampling of households from the area frame 13
5.4.2 Sampling of households from the list frame of telephone numbers 15
5.4.3 Sampling of households from the Random Digit Dialing frame of telephone numbers 15
5.5 Sample allocation over the collection periods 16
5.6 Sampling of interviewees 16
5.7 Supplementary buy-in sample in Ontario 17

6. DATA COLLECTION 19

6.1 Computer-assisted interviewing 19
6.2 CCHS application development 19
6.3 Interviewer training 20
6.4 The interview 21
6.5 Field operations 23
6.6 Quality control and collection management 24

7. DATA PROCESSING 25

7.1 Editing 25
7.2 Coding 25
7.3 Creation of derived variables 25
7.4 Weighting 25
7.5 Income imputation 26

8. WEIGHTING 27

8.1 Overview 27
8.2 Weighting of the area frame sample 28
8.3 Weighting of the telephone frame sample 29
8.4 Integration of the telephone and area frames (I1) 31
8.5 Post-integration weighting steps 31
8.6 Particular aspects of the weighting in the three territories 33
8.7 Creation of a share weight 34
8.8 Weighting for a two-year file 34

9. DATA QUALITY 35

9.1 Response rates for 2012 35
9.2 Response rates for 2011-2012 35

9.3 Survey Errors 36
9.3.1 Non-sampling Errors *37*
9.3.2 Sampling Errors *37*

10. GUIDELINES FOR TABULATION, ANALYSIS AND RELEASE 38

10.1 Rounding guidelines 38
10.2 Sample weighting guidelines for tabulation 39
10.2.1 Definitions: categorical estimates, quantitative estimates *39*
10.2.2 Tabulation of categorical estimates *40*
10.2.3 Tabulation of quantitative estimates *40*
10.3 Guidelines for statistical analysis 40
10.4 Release guidelines 41

11. APPROXIMATE SAMPLING VARIABILITY TABLES 43

11.1 How to use the CV tables for categorical estimates 43
11.2 Examples of using the CV tables for categorical estimates 45
11.3 How to use the CV tables to obtain confidence limits 48
11.4 Example of using the CV tables to obtain confidence limits 49
11.5 How to use the CV tables to do a Z-test 49
11.6 Example of using the CV tables to do a Z-test 49
11.7 Exact variances/coefficients of variation 50
11.8 Release cut-offs for the CCHS 51

12. MICRODATA FILES: DESCRIPTION, ACCESS AND USE 52

12.1 Master files 52
12.1.1 Research Data Centre *52*
12.1.2 Custom tabulations *52*
12.1.3 Remote access *52*
12.2 Share files 52
12.3 Public use microdata files 53
12.4 How to use the CCHS data files: annual data file or two-year data file? 54
12.5 Use of weight variable 55
12.6 Variable naming convention beginning in 2007 55
12.7 Variable naming convention before 2007 57
12.8 Guidelines for the use of sub-sample variables – Not applicable to 2012 and 2011-2012 data files 57
12.9 Data dictionaries 57
12.10 Differences in calculation of common content variables using different files 58
Appendix A – Canadian community health survey content overview (2011-2012) 59
Appendix B – Selection of optional content by province and territory (2012 and 2011-2012) 62
Appendix C - Available geography in the master and share files and their corresponding codes: Canada, provinces/territories, health regions and peer groups 69
Appendix D (2012) Sample allocation by health region and frame, and sample allocation by Local Health Integrated Network (LHIN) and frame in Ontario 74
Appendix E (2012) - Response rates by health region and frame, and response rates by Local Health Integrated Network (LHIN) and frame in Ontario 79
Appendix F (2011-2012) - Sample allocation by health region and frame and sample allocation by Local Health Integrated Network (LHIN) and frame in Ontario 86
Appendix G (2011-2012) - Response rates by health region and frame and response rates by Local Health Integrated Network (LHIN) and frame in Ontario 91

WHAT'S NEW IN THE 2012 CANADIAN COMMUNITY HEALTH SURVEY?

Content

The following modifications were made to the Canadian Community Health Survey (CCHS) – Annual component questionnaire in 2011 and 2012:

Changes

- **Restriction of Activities (RAC)**: This module has been included in the core component of the questionnaire throughout the years, with the exception of 2011 when it was only included in the territories. The module was suspended in 2011 to avoid overlapping content with two disability modules that were collected as rapid response: Adult Disability Identifiers – ADI and Disability Screening Questions – DSQ.
- **Consultations about Mental Health (CMH):** This module was previously offered as an optional module. In 2011 – 2012, it was moved as part of the 2 year common content on mental well being (see Appendix A for details).
- **Contact with Health Professionals (CHP)**: In 2012, this module was split into two modules. The first series of questions became part of core content (CHP) and includes questions on consultations with family doctor, eye specialist, other specialists, nurse and dentist in the past 12 months. The second set of questions part was moved in the optional component (Consultations with Health Professionals – Part 2 [CP2]).

New modules

- **Positive Mental Health (PMH)**: This new module was introduced in 2011 as part of the 2 year common content on mental well being.
- **Physical activities – Stages of Changes (SCP)**: This new module was added as an optional module in 2011.
- **Social Provisions (SPS)**: This new module was also added in the optional component in 2011. It includes questions on five social provisions sub-scales - Attachment, Social integration, Reassurance of worth, Reliable alliance, and Guidance.
- **Contact with Health Professionals – Part 2 (CP2):** This optional module was previously part of one single Consultation with Health Professional (CHP) before 2012. It includes questions on consultations with chiropractors, physiotherapists, psychologists, social workers and audiologists in the past 12 months.

Geography

- As a result of changes to Health Region geography in Manitoba in 2012, data from 2012 onward will be presented by Regional Health Authority (RHA). The 11 Health Regions in Manitoba were merged into 5 RHA's as follows:
  - Winnipeg RHA (4610 A) and Churchill RHA (4690 F) were merged to form Winnipeg RHA (4601 A)
  - Assiniboine RHA (4645 D), Parkland (4660 D) and Brandon RHA (4615 A) were merged to form Prairie Mountain Health (4602 D)
  - Interlake RHA (4630 E) and North Eastman RHA (4620 E) were merged to form Interlake-Eastern RHA (4603 E)
  - NOR-MAN RHA (4670 H) and Burntwood RHA (4680 F) were merged to form Northern RHA (4604 F)
  - Central RHA (4640 D) and South Eastman RHA (4625 E) were merged to form Southern Health (4605 D)

Data Processing

- Beginning with the 2011 reference year, the household income variable will be imputed. Missing values due to either respondent refusal or respondent's lack of knowledge of household income will be completed using statistical techniques. The main variable of interest is INC_3: 'Total household income - best estimate' but all variables that are derived based on income will also be affected. The income variables along with an imputation flag (INCFIMP) indicating which values were imputed will be provided on the data file. For more information on the imputation process, please refer to the document 'Income Imputation for the Canadian Community Health Survey' available under the 'Documentation' section of the Definitions, data sources and methods page on the Statistics Canada website.
- In 2012, this variable was renamed to INCFIMP4. New categories were created to make it more detailed.

Documentation

- **Errata:** Errors that are discovered in the CCHS data files and products after dissemination are communicated to users in a cumulative error log called "CCHS Errata". The errata provides details on:
  - the products affected (e.g. master or share files or PUMF);
  - the years affected;
  - suggested corrections (if possible); and
  - steps to carry out to apply the correction

  Since the release of the 2011 CCHS, there have been three additional entries to the errata document. The subjects of these entries are:
  - Labeling errors in HMC_13
  - Labeling errors in ALW_2A1, ALW_2A2, ALW_2A3, ALW_2A4, ALW_2A5, ALW_2A6, ALW_2A7
  - Labeling errors in SDC_4, question on ancestry

  An up to date copy of the CCHS Errata can be obtained by contacting Health Statistics Division at 613-951-1746 or hd-ds@statcan.gc.ca.

1. INTRODUCTION

The Canadian Community Health Survey (CCHS) is a cross-sectional survey that collects information related to health status, health care utilization and health determinants for the Canadian population. It surveys a large sample of respondents and is designed to provide reliable estimates at the health region level. In 2007, major changes were made to the CCHS design. Data is now collected on an ongoing basis with annual releases, rather than every two years as was the case prior to 2007. The survey's objectives were also revised and are as follows:

- support health surveillance programs by providing health data at the national, provincial and intra-provincial levels;
- provide a single data source for health research on small populations and rare characteristics;
- timely release of information easily accessible to a diverse community of users; and
- create a flexible survey instrument that includes a rapid response option to address emerging issues related to the health of the population.

Details of the other redesign changes are provided in section 3.

The CCHS data is always collected from persons aged 12 and over living in private dwellings in the 115 health regions covering all provinces and territories. Excluded from the sampling frame are individuals living on Indian Reserves and on Crown Lands, institutional residents, full-time members of the Canadian Forces, and residents of certain remote regions. The CCHS covers approximately 98% of the Canadian population aged 12 and over.

The purpose of this document is to facilitate the manipulation of the CCHS microdata files and to describe the methodology used. The CCHS produces three types of microdata files: master files, share files and public use microdata files (PUMF). The characteristics of each of these files are presented in this guide. The PUMF is released every two years and contains two years of data. The next PUMF file will be released in September 2013 and will include the data collected for the years 2011and 2012.
Any questions about the data sets or their use should be directed to:

Electronic Products Help Line: 1-800-949-9491

For custom tabulations or general data support:
Client Custom Services, Health Statistics Division: 613-951-1746
E-mail: hd-ds@statcan.gc.ca

For remote access support: 613-951-1746
E-mail: cchs-escc@statcan.gc.ca
Fax: 613-951-0792

2. BACKGROUND

In 1991, the National Task Force on Health Information cited a number of issues and problems with the health information system. The members felt that data was fragmented; incomplete, could not be easily shared, was not being analysed to the fullest extent, and the results of research were not consistently reaching Canadians.[1]

In responding to these issues, the Canadian Institute for Health Information (CIHI), Statistics Canada and Health Canada joined forces to create a Health Information Roadmap. From this mandate, the Canadian Community Health Survey (CCHS) was conceived. The format, content and objectives of the CCHS evolved through extensive consultation with key experts and federal, provincial and community health region stakeholders to determine their data requirements.[2]

To meet many data requirements, the CCHS had a two-year data collection cycle. Until the redesign in 2007, the first year of the survey cycle, designated by ".1", was a general population health survey, designed to provide reliable estimates at the health region level. The second year of the survey cycle, designated by ".2", had a smaller sample and was designed to provide provincial level results on specific health topics.

New designations for Cycles .1 and .2

As of 2007, the regional component of the CCHS program began being collected on an ongoing basis. To avoid confusion with the health focused surveys, the two components stopped using the ".1" and ".2" designations to distinguish them. Henceforth, the x.1 cycles of the CCHS are designated as "the annual component" of the CCHS. The full title is "The Canadian Community Health Survey – Annual component, 2011" and the short title is simply "CCHS –2011". The focused content component of the survey remains unchanged. It will continue to examine in greater detail more specific topics or populations. It will be designated by the name of the survey followed by the topic of the themes covered by each survey (e.g., "Canadian Community Health Survey on Healthy Aging" or "CCHS – Healthy Aging").

[1] 1999. Health Information Roadmap: Responding to Needs, Health Canada, Statistics Canada. p. 3.

[2] 1999. Health Information Roadmap: Beginning the Journey. Canadian Institute for Health Information/Statistics Canada. ISBN 1-895581-70-2. p. 19.

3. CCHS REDESIGN IN 2007

Until 2005, the CCHS data were collected every two years over a one year period and released every two years, about six months after the end of the collection period. There were two main objectives for the 2007 CCHS redesign: to address the needs of partners to increase the survey's content and the frequency of data releases, and to ensure better use of operational resources. For these reasons, the proposed changes to the CCHS design focused on improving the survey's efficiency and flexibility through ongoing data collection.

Extensive consultations were held across Canada with key experts and federal, provincial and health region stakeholders to gather input on the proposed changes and detailed information on the data requirements and products of the various partners.

Below are the main changes arising from the CCHS redesign:

- In the past, the CCHS data were collected from 130,000 respondents over a 12-month period. Now, data collection takes place on an ongoing basis. The sample, which retains the same size, is divided into 12 two-month collection periods. Each collection period is representative of the population living in the ten Canadian provinces during the two months. For operational reasons, the sample in the territories is representative of their population after 12 months.

- The common content component is divided into three: the annual common content (previously referred to as core content), the one year and two-year common content (previously referred to as theme content). The one year common content is asked for one year and re-introduced every two or four years. The two year common content is asked for two years and re-introduced every four years. The two year and one year common content was created to take advantage of the continuous collection approach. The data collection time for this component can be adjusted based on the prevalence of the desired estimates and their geographic level. The annual common content will remain relatively stable over time. At the discretion of the provinces and regions, the optional content can also be adjusted on an annual basis, rather than every two years.

- Content and collection changes inevitably impact the dissemination strategy. Previously, data were released every two years. Since 2008, CCHS data are released annually. Every two years, a file combining the two years' sample (130,000 respondents) is also released. In addition to these regular files, other special files will be made available when additional content has been collected during collection periods that do not correspond to the standard annual periods, which is January to December.

- The annual data collection is divided into six two-month periods. Unlike the previous collection strategy, these periods no longer overlap. This provides more efficient oversight of collection and offers the possibility of changing the collection interface every two months, if necessary.

4. CONTENT STRUCTURE OF THE CCHS

In addition to socio-demographic and administrative data, the content of the CCHS includes three components, each of which addresses a different need: the common content component comprising the annual common content, the two year and one year common content, the optional content component, and the rapid response component. Appendix A lists the modules included in the 2011 and 2012 questionnaire by component.

The average length of a CCHS interview is estimated at 40 to 45 minutes.

Table 4.1 Length of survey by component

CCHS component	Average interview time
Common content • Annual • One and two-year	**30 minutes** (20 minutes) (10 minutes)
Optional content	**10 minutes**
Rapid response content (As requested on a cost recovery basis)	**2 minutes**

4.1 Common content

The CCHS common content component includes questions asked of respondents in all provinces and territories (unless otherwise specified). It is divided into three components: the annual common content, one-year and two year common content.

The annual common content consists of questions asked of all survey respondents. These questions will remain relatively stable in the questionnaire for a period of about six years up to 2014, unless a major concern is raised about quality.

The one year and two-year common content (previously called theme content) comprises questions related to a specific topic. Combined, the two-year and one year common content take about 10 minutes of the interview time. Modules comprising this content type could be reintroduced in the survey every two, four or six years, if required. This component enables CCHS to better plan its content in the medium term.

Some of the modules in the one year common content may be asked of a sub-sample of respondents if the objective of these questions is to provide reliable data at the national or provincial level, rather than at the health region level. This approach is used to minimize the related response burden and costs.

4.2 Optional content

The optional content component gives health regions the opportunity to select content that addresses their provincial or regional public health priorities. The optional content is selected from a long list of modules available for inclusion in the CCHS. The content modules selected by a region are asked only of residents in the regions that selected these modules. In reality, since 2005 (cycle 3.1), the regions and provinces have opted to coordinate the optional content selected in order to ensure a uniform selection of optional modules provincially. The optional content may vary annually depending on needs and must be reviewed every two years.

It should be noted that, unlike the modules included in the common content, the resulting data from the optional content modules is not easily generalized across Canada[3].

Appendix B presents the selection results of the optional content for the current year and two year combined by province of residence.

4.3 Rapid response content

The rapid response component is offered on a cost-recovery basis to organizations interested in obtaining national estimates on an emerging or specific topic related to the health of the population. The rapid response content takes a maximum of two minutes of interview time. The questions appear in the questionnaire for a single collection period (two months) and are asked of all CCHS respondents during that period, excluding the Territories.

4.4 Content included in data files

The survey produces different data files:

- one year reference period
- combined two years reference periods and
- one year sub-sample data files.

Table 4.2 provides clarification about the data files available for the 2011 and 2012 CCHS.

One year data files

The survey produces data files every year. In June 2013, an annual file based on the 2012 reference period has been released. It includes respondents from the 2012 data collection and variables from the common annual content, common one year content, common two year content as well as optional content.

[3] Unless all health regions in Canada select an optional module in the same collection period, which has never happened to date.

Two year data files

Every two years, a file combining the most recent two years is released. A combined file also to be released in June 2013 will contain data from 2011 and 2012. The following two year file is scheduled to be released in 2015, and will include both the 2013 and 2014 reference years.

The two-year data file includes all respondents and the questions that were in the survey over the two year reference period. Unless otherwise specified, it is the question component from the common annual and two-year content and selected optional content over the two year period. The one-year common content and optional content selected for one year only are not available in the two-year data file.

Sub-sample data files

Any modules collected from a sub-sample of the population will continue to be disseminated in separate files. These files include the annual and one year common content collected from a sub-sample of respondents. Sub-sample files have been released as follow:

Year	**Modules**
2000	Waiting times and Access to health care services
2003	Dental visits, Driving and safety, Health utility index, Medication use, Oral health 2
2005	Waiting times, Access to health care services, Patient satisfaction, Health Utility Index, Measured height and weight, Fruit and vegetable consumption, Labour force – long form
2007	Waiting times, Access to health care services and Patient satisfaction
2008	Measured height and weight
2009	Waiting times, Access to health care services
2011	Waiting times, Access to health care services

Table 4.2 Content components for the 2011 and 2012 data files

Files		Annual common content	2011 one year common content[1]	2012 one year common content	2011-2012 two-year common content	Optional content[2]
2011	**Main**	Yes	No	N/A	Yes	Yes
	Sub-sample (2 modules)	Yes	Yes	N/A	No	No
2012	**Main**	Yes	N/A	Yes	Yes	Yes
2011-2012	**Main**	Yes	No	No	Yes	Yes

1 The 2011 annual common content was comprised of two modules (Access to health care services and Waiting times) which were all asked to a sub-sample of respondents.

2 Optional content will be included in the 2011-2012 data file if it is asked of respondents in a jurisdiction during the two year period. Otherwise, it will only be included in the file of the year in which it was collected. Note that if an annual common content module from one year is selected for the optional content of a jurisdiction during the second year, the module will be included in the two-year data file and will be processed as optional content.

5. SAMPLE DESIGN

5.1 Target population

The CCHS targets persons aged 12 years and older who are living in private dwellings in the ten provinces and three territories. Persons living on Indian Reserves or Crown lands, those residing in institutions, full-time members of the Canadian Forces and residents of certain remote regions are excluded from this survey. The CCHS covers approximately 98% of the Canadian population aged 12 and older.

5.2 Health regions

For administrative purposes, each province is divided into health regions (HR) and each territory is designated as a single HR. Statistics Canada is sometimes asked to make minor changes to the boundaries of some of the HRs to correspond to the geography of the Census, or to better account for the health data needs determined by the new geographic boundaries. For CCHS 2012, data was collected in 107 HRs in the ten provinces, as well as to one HR per territory, totalling 110 HRs (Appendix C).

The definition of HRs was modified between the time of sampling and the creation of the data files for two provinces. In Prince Edward Island, the 3 HRs defined at the time of sampling (1101, 1102 and 1103) have been aggregated into a single HR (1100). In Manitoba, the 11 Regional Health Authorities (RHA) that were forming 10 HRs at the time of sampling have been grouped to form 5 new HRs[4]. The current chapter on sample design, as well as the figures on sample sizes provided in Appendix D and Appendix F, refer to the HRs as they were defined at the time of sampling.

5.3 Sample size and allocation

To provide reliable estimates for each HR given the budget allocated to the CCHS component, it was determined that the survey should consist of a sample of nearly 130,000 respondents over a period of 2 years. Although producing reliable estimates for each HR was a primary objective, the quality of the estimates for certain key characteristics at the provincial level was also deemed important. Therefore, the sample allocation strategy, consisting of three steps, gave relatively equal importance to the HRs and the provinces. In the first step, a minimum size of 500 respondents per HR was imposed. This is considered the minimum for obtaining a reasonable level of data quality. However, due to response burden, a maximum sampling fraction of 1 out of 20 dwellings was imposed to avoid sampling too many dwellings in smaller regions also targeted by other surveys. Note that very few HRs have a size lower than 500 due to limit of the sampling fraction. In this first step, 60,350 units were allocated in total. The second step involves allocating the rest of the available sample by using an allocation proportional to the population size by province. The total sample size by province is therefore the sum of the sizes established by the two first steps. This

[4] The 11 RHAs were corresponding to the HRs used at the time of sampling, except for the RHA 4680 (Burntwood) and 4690 (Churchill), which were grouped together to form the HR 4685 for the CCHS. The correspondence between the 5 new HRs and these 11 RHAs is as follows: HR 4601 (4610 and 4690), HR 4602 (4615, 4645 and 4660), HR 4603 (4620 and 4630), HR 4604 (4670 and 4680) and HR 4605 (4625 and 4640).

sample allocation strategy was used for CCHS 2005 and the sample sizes have remained mainly the same since then. The sample was then divided evenly between the 2 collection years. Table 5.1 gives the targeted sample sizes for 2012 and for 2011-2012.

Table 5.1 Number of health regions and targeted sample sizes by province/territory, 2012 and 2011-2012

Province	Number of HRs	Targeted sample size 2012	Targeted sample size 2011-2012
Newfoundland and Labrador	4	2,005	4,010
Prince Edward Island	3	1,001	2,002
Nova Scotia	6	2,521	5,042
New Brunswick	7	2,575	5,150
Quebec	16	12,146	24,291
Ontario[1]	36	22,198	44,396
Manitoba	10	3,750	7,500
Saskatchewan	11	3,860	7,720
Alberta	5	6,097	12,197
British Columbia	16	8,045	16,090
Yukon	1	600	1,200
Northwest Territories	1	600	1,200
Nunavut	1	350	700
Canada	**117**	**65,748**	**131,498**

[1] The sample size for Ontario includes the buy-in extra sample by LHIN. The initial annual sample size for Ontario before the buy-in was 20,880 units (refer to section 5.7 for further details).

In the third step, the provincial sample was allocated among its HRs proportionally to the square root of the estimated population in each HR[5]. This three-step approach gives sufficient sample for each HR with minimal disturbance to the proportionality of the allocation by province.

Note that the three territories were not part of the above allocation strategy as they were dealt with separately. Each year, 600 sample units were allocated to the Yukon, 600 to the Northwest Territories and 350 to Nunavut. These sizes are determined according to the available budget. The sample allocation for the territories is done proportionally to the population sizes of the strata. The strata used were the same as those defined by the Labour Force Survey (LFS), which group together communities (for more details, see section 5.4.1).

The sample was then divided between the area frame and the list frame[6], as described in the next section. It should be mentioned that the number of units taken from each frame was increased in order to account for the anticipated out-of-scope and non-response rates based on the rates obtained

[5] The allocation of the provincial sample among its HRs is based on the geographical boundaries and population counts of the HRs at the time of the 2007 redesign.

[6] Except for two regions which use a random digit dialing frame (RDD) only (section 5.4.3) and the three territories which use only area frame and random digit dialing frame (RDD) (sections 5.4.1 and 5.4.3).

in previous CCHS cycles. These adjusted sample sizes are called the *raw sample sizes*, and correspond to the number of units that have to be selected in order to obtain the targeted sample sizes. The sample sizes by HR and frame are provided in Appendix D for 2012 and in Appendix F for 2011-2012.

5.4 Frames, household sampling strategies

The CCHS used three sampling frames to select the sample of households: 40.5% of the sample of households came from an area frame, 58.5% came from a list frame of telephone numbers and the remaining 1% came from a Random Digit Dialling (RDD) sampling frame.

5.4.1 Sampling of households from the area frame

The CCHS used the area frame designed for the Canadian Labour Force Survey (LFS) as a sampling frame. The sampling plan of the LFS is a multistage stratified cluster design in which the dwelling is the final sampling unit[7]. In the first stage, homogeneous strata are formed and independent samples of clusters are drawn from each stratum. In the second stage, dwelling lists are prepared for each cluster and dwellings, or households, are selected from these lists.

For the purpose of the LFS plan, each province is divided into three types of regions: major urban centres, cities, and rural regions. Geographic or socio-economic strata are created within each major urban centre. Within the strata, between 150 and 250 dwellings are grouped together to create clusters. Some urban centres have separate strata for apartments or for census Dissemination Areas (DA) to pinpoint households with high income, immigrants and aboriginals. In each stratum, six clusters or residential buildings (sometimes 12 or 18 apartments) are chosen by a random sampling method with a probability proportional to size (PPS), the size of which corresponds to the number of households. The number six is used throughout the sample design to allow for one sixth of the LFS sample to be rotated each month.

The other cities and rural regions of each province are stratified first on a geographical basis, then according to socio-economic characteristics. In the majority of strata, six clusters (usually census DAs) are selected using the PPS method. Some geographically isolated urban centres are covered by a three-stage sampling design. This type of sampling plan is used for Quebec, Ontario, Alberta and British Columbia.

Once the new clusters are listed, the sample is obtained using a systematic sampling of dwellings. The sample size for each systematic sample is called the "yield". Table 5.2 gives an overview of the types of PSUs used in the LFS sample and the yield predicted by systematic sample. As the sampling rates are determined in advance, there is frequently a difference between the expected sample size and the numbers that are obtained. The yield of the sample, for example, is sometimes excessive. This can particularly happen in sectors where there is an increase in the number of dwellings due to new construction. To reduce the cost of collection, an excessive output is corrected

[7] Statistics Canada (2008). *Methodology of the Canadian Labour Force Survey.* Statistics Canada. Cat. No. 71-526-XIE.

by eliminating, from the beginning, a part of the units selected and by modifying the weight of the sample design. This change is dealt with during weighting.

Table 5.2 Major first-stage units, sizes and yields

Area	Primary Sampling Unit (PSU)	Size (households per PSU)	Yield (sampled households)
Toronto, Montreal, Vancouver	Cluster	150-250	6
Other cities	Cluster	150-250	8
Most rural areas / small urban centres	Cluster	100-250	10

Due to the specific of the CCHS, some modifications had to be incorporated in this sampling strategy. To obtain an annual sample of about 26,000 respondents for a given year of CCHS, about 40,000 dwellings had to be selected from the area frame to account for vacant dwellings and non-responding households. Each month, the LFS design provides approximately 60,000 dwellings distributed across the various economic regions in the ten provinces, whereas the CCHS required 40,000 dwellings distributed across the HRs, which have different geographic boundaries from those of the LFS economic regions. Overall, the CCHS required a lower number of dwellings than those generated by the LFS selection mechanism, which corresponds to an average *adjustment factor* of 0.66 (40,000/60,000). . However, since the adjustment factors varied at the HR level, certain adjustments were required.

The changes made to the selection mechanism in the regions varied depending on the size of the adjustment factors. For HRs that had a factor smaller than or equal to 1, the number of PSUs selected was reduced if necessary. For example, if the factor was 0.5 then only 3 PSUs were selected in each stratum instead of the usual number of 6 PSUs. For those HRs with a factor greater than 1 but smaller than or equal to 2, the sampling process of dwellings within a PSU was repeated for a subset of the selected PSUs that were part of the same HR. For example, if the factor was 1.6 then the selection of dwellings within a PSU was repeated for 4 of the 6 PSUs in all strata of that HR. When it was necessary to have a repeated selection of dwellings within a PSU and there were no more dwellings available in that PSU, then another PSU was selected. When the factor was greater than 2, the sampling process of dwellings was repeated among other PSUs that were part of the same HR[8].

Finally, when the number of dwellings available in the selected PSUs was greater than the requested number of dwellings for a given HR, a sub-sample of dwellings was selected. This process is called *stabilization*.

[8] To reduce listing costs, the sampling process of dwellings was repeated up to 3 times within PSUs already selected in urban areas only. These cases were exceptions, however.

Sampling of households from the area frame in the three territories

For operational reasons, the LFS area frame sample design for the three territories was different. For each territory, the larger communities each have their own stratum while smaller communities are grouped into strata based on various characteristics (population, geographical information, proportion of Inuit and/or Aboriginal persons, and median household income). The LFS defined five design strata in the Yukon, ten in the Northwest Territories and seven in Nunavut. For strata consisting of a group of communities, the first stage of selection consisted of randomly selecting one community with a probability proportional to population size within each design stratum. Then, within the selected community, the second stage consisted of selecting households using the same sampling strategy as the one described above. The CCHS selected its sample from the same communities sampled by the LFS, while ensuring that different dwellings were selected. If too many or too few dwellings were available for a community within a stratum, another community was selected for the CCHS. For larger communities with their own stratum, only one stage design was necessary where households were selected directly using the same sampling strategy described above.

It is worth mentioning that the frame for the CCHS covered 90% of the private households in the Yukon, 97% in the Northwest Territories and 71% in Nunavut[9].

5.4.2 Sampling of households from the list frame of telephone numbers

With the exception of 5 HRs (the two RDD-only HRs and the three territories), the list frame of telephone numbers was used in all HRs to complement the area frame. The list frame consists of the Canada Phone directory which is an external administrative database of names, addresses and telephone numbers from telephone directories in Canada updated every six months. It was linked to administrative postal code conversion files to map each telephone number to a stratum. Within each stratum, the required number of telephone numbers was selected using a simple random sampling process from the list. As for the RDD frame, additional telephone numbers were selected to account for the numbers not in service or out-of-scope.

It is important to mention that the undercoverage of the list frame is higher than the one for the RDD as unlisted numbers do not have a chance of being selected. Nevertheless, as the list frame is always used as a complement to the area frame, the impact of the undercoverage of the list frame is minimal and is dealt with during weighting.

5.4.3 Sampling of households from the Random Digit Dialing frame of telephone numbers

In four HRs, a Random Digit Dialling (RDD) sampling frame of telephone numbers was used to select a sample of households. The sampling of households from the RDD frame used the Elimination of Non-Working Banks (ENWB) method, a procedure adopted by the General Social

[9] In Nunavut, only the 10 largest communities are covered by the survey: Iqaluit, Cambridge Bay, Baker Lake, Arviat, Rankin Inlet, Kugluktuk, Pond Inlet, Cape Dorset, Pangnirtung and Igloolik.

Survey[10]. A bank of one hundred telephone numbers (the first eight digits of a ten-digit telephone number) is considered to be non-working if it does not contain any residential telephone numbers. At first, the frame consists of a list of all possible banks and, as non-working banks are identified, they are eliminated from the frame. It should be noted that these banks are eliminated only when there is evidence from various sources that they are non-working. When there is no information about a bank it is left on the frame. The Canada Phone Directory and telephone companies' billing address files were used in conjunction with various internal administrative files to eliminate non-working banks.

Using available geographic information (postal codes), the banks on the frame were regrouped to create RDD strata to encompass, as closely as possible, the HR areas. Within each RDD stratum, a bank was randomly chosen and a number between 00 and 99 was generated at random to create a complete, ten-digit telephone number. This procedure was repeated until the required number of telephone numbers within the RDD stratum was reached. Frequently, the number generated is not in service or is out-of-scope, and therefore, many additional numbers must be generated to reach the targeted sample size. This success rate varies from region to region. Within the CCHS, the success rates ranged from 25% to 50% among the four HRs which required the use of the RDD frame.

5.5 Sample allocation over the collection periods

In order to balance interviewer workload and to minimize possible seasonal effects on estimates, the initial sample size for each frame is allocated equally over the six 2-month collection periods. Exceptionally in 2011, this was slightly modified because of the additional workload related to the collection of data for the 2011 Census, which started in May. A portion of the CCHS sample allocated to the May-June and July-August collection periods was transferred to the other four collection periods.

In the area frame, each PSU selected within each HR was randomly assigned to a collection period accounting for a number of constraints related to field operations or weighting, while maintaining a uniform size for each period. For example, a sample that is representative of the Canadian population is ensured every six months by ensuring that the dwelling sample covers all LFS strata during this period.

For the lists of telephone numbers, independent samples were selected in each collection period. This strategy ensures that each sample is representative of the in scope Canadian population for each two month period.

5.6 Sampling of interviewees

As was done for the previous cycles, the selection of individual respondents was designed to ensure over-representation of youths (12 to 19). The selection strategy that was adopted accounted for user needs, cost, design efficiency, response burden and operational constraints. One person is selected per household using varying probabilities taking into account the age and the household

[10] Norris, D.A. and Paton, D.G. (1991). Canada's General Social Survey: Five Years of Experience, *Survey Methodology*, 17, 227-240.

composition. The selection probabilities resulted from simulations using various parameters in order to determine the optimal approach without causing extreme sampling weights.

Table 5.3 gives the selection weight multiplicative factors used to determine the probabilities of selection of individuals in sampled households by age group. For example, for a three-person household formed of two adults of age 45 to 64 and one 15-year-old, the teenager would have a 7/9 chance of being selected (i.e., 70/(70+10+10)) while each of the adults would have a 1/9 chance of being selected. To avoid extreme sampling weights, there is one exception to this rule: if the size of the household is greater than or equal to 5 or if the number of 12-19 year olds is greater than or equal to 3 then the selection weight multiplicative factor equals 1 for each individual in the household. Consequently, all people in that household have the same probability of being selected.

Table 5.3 Selection weight multiplicative factors for the person-level sampling strategy by age

Selection Weight Multiplicative Factors					
Age	**12-19**	**20-29**	**30-44**	**45-64**	**65+**
Factors	70	50	20	10	10

5.7 Supplementary buy-in sample in Ontario

The province of Ontario requested a sample increase in order to produce estimates at the Local Health Integrated Network (LHIN) geography level. Ontario contains 14 LHIN. The CCHS sample was increased in order to obtain a minimum size of 2,000 respondents per LHIN over a period of 2 years. As the HR and LHIN boundaries intersect each other, the stratification level used was the HR-LHIN overlap. The preliminary sample sizes allotted by HR are therefore preserved. In cases where the HR allocation prevented the sample from reaching sizes of 2,000 respondents per LHIN, the sample was then increased, and was allocated proportionally to the size of the population within the HR-LHIN overlap. Table 5.4 provides the targeted sample sizes by LHIN for 2012 and 2011-2012.

Table 5.4 Targeted sample sizes by Local Health Integrated Network (LHIN), CCHS 2012 and 2011-2012

LHIN	Targeted sample size	Targeted sample size 2011-2012
01-Erie St. Clair	1,550	3,100
02-South West	2,561	5,122
03-Waterloo Wellington	1,243	2,486
04-Hamilton Niagara Haldimand Brant	2,597	5,194
05-Central West	1,049	2,098
06-Mississauga Halton	1,115	2,230
07-Toronto Central	1,123	2,246
08-Central	1,411	2,822
09-Central East	2,108	4,216
10-South East	1,313	2,626
11-Champlain	2,057	4,114
12-North Simcoe Muskoka	1,041	2,082
13-North East	1,990	3,980
14-North West	1,040	2,080
Ontario	**22,198**	**44,396**

The total sample size of the HR-LHIN overlapping areas was then allocated between the two frames, with 59% going to the list frame and the 41% to area frame. The usual sample selection procedures within each frame were then applied to the total sample. The additional sample was included as part of the full CCHS sample. Sample sizes by Local Health Integrated Network and frame are given in Appendix D for 2012 and in Appendix F for 2011-2012.

6. DATA COLLECTION

6.1 Computer-assisted interviewing

Between January and December 2012, over 60,000 valid interviews were conducted using computer assisted interviewing (CAI). Approximately 40% of the interviews were conducted in person using computer assisted personal interviewing (CAPI) and the other 60% were conducted over the phone using computer assisted telephone interviewing (CATI).

CAI offers two main advantages over other collection methods. First, CAI offers a case management system and data transmission functionality. The case management system automatically records important management information during each attempt on a case and provides collection management reports. Additionally, the case management system routes the questionnaire applications and sample files from Statistics Canada's main office to regional collection offices (in the case of CATI) and from the regional offices to the interviewers laptops (for CAPI). Data returning to the main office take the reverse route. To ensure confidentiality, the data are encrypted before transmission. The data are then unencrypted once they are on a separate secure computer with no remote access. CAI also provides an automated call scheduler, i.e. a central system to optimise the timing of call-backs and the scheduling of appointments used to support CATI collection.

Second, CAI allows for custom interviews for every respondent based on their individual characteristics and survey responses. This includes:

- automatic skips of questions that are not applicable to the respondent
- edits to check for inconsistent answers or out-of-range responses are applied automatically and on-screen prompts are shown when an invalid entry is recorded. Immediate feedback is given to the respondent and the interviewer is able to correct any inconsistencies.
- question text, including reference periods and pronouns, is customised automatically based on factors such as the age and sex of the respondent, the date of the interview and answers to previous questions.

6.2 CCHS application development

The CCHS uses two separate CAI applications to collect data: one for telephone interviews (CATI) and one for personal interviews (CAPI). This is done in order to customise each application's functionality to the type of interview being conducted. Each application consists of entry, health content, and an exit component.

Entry and exit components contain standard sets of questions designed to guide the interviewer through contact initiation, respondent selection, collection of important sample information and determination of cases status. The health content consists of the health modules and makes up the bulk of the application. This includes common modules asked of all respondents and optional

modules which differ by health region. Each application underwent three stages of testing: block, integrated and end to end.

Block level testing consists of independently testing each content module or "block" to ensure skip patterns, logic flows and text, in both official languages, are specified correctly. Skip patterns or logic flows across modules are not tested at this stage as each module is treated as a standalone questionnaire. Once all blocks are verified by several testers they are added together along with entry and exit components into integrated applications. These newly integrated applications are then ready for the next stage of testing.

Integrated testing occurs when all of the tested modules are compiled with the entry and exit components, into an integrated application. This second stage of testing ensures that key information such as age and gender are passed from the entry to the health content and exit components of the applications. It also ensures that variables affecting skip patterns and logic flows are correctly passed between modules within the health content. At this stage, all possible scenarios faced by interviewers can be simulated to ensure proper functionality. Since the applications essentially function as they will in the field, these scenarios are able to test various aspects of the entry and exit components, including establishing contact, collecting contact information, determining whether a case is in scope, rostering households, creating appointments and selecting respondents. The applications are also tested to ensure that during an interview, correct modules are triggered reflecting health region optional content selections.

End to end testing occurs when the fully integrated applications are placed in a simulated collection environment. The applications are loaded onto computers that are connected to a test server. Data are then collected, transmitted and extracted in real time, exactly as it would be done in the field. This last stage allows for the testing of all technical aspects of data input, transmission and extraction for each of the CCHS applications. It also provides a final opportunity to find errors within the entry, health content and exit components.

6.3 Interviewer training

Project managers, senior interviewers and interviewers from regional collection offices were sent self study training packages before the start of collection. These packages were prepared by the CCHS project team and were used by existing experienced CCHS interviewers to reinforce their previous training. Project managers and senior interviewers also conducted customised training sessions for new CCHS interviewing staff as needed. Additionally, training sessions were conducted on a monthly basis in order to get interviewers comfortable using the CCHS 2012 applications, familiarise them with survey content and to introduce them to interviewing procedures specific to the CCHS. The training focused on:

- goals and objectives of the survey including a focus on the survey redesign
- survey methodology
- application functionality

- reviewing the questionnaire content and completing exercises that emphasized significant content changes
- interviewer techniques for maintaining response and completing exercises to minimise non-response
- use of mock interviews to simulate difficult situations and practise potential non-response situations
- survey management
- transmission procedures

One of the key aspects of the training was a focus on minimizing non-response. Exercises to minimise non-response were prepared for interviewers. The purpose of these exercises was to have the interviewers practice convincing reluctant respondents to participate in the survey. There was also a series of refusal avoidance workshops given to the senior interviewers responsible for refusal conversion in each regional collection office.

6.4 The interview

Sample units selected from the telephone list and RDD (Random Digit Dialling) frames were interviewed from centralised call centres using CATI. The CATI interviewers were supervised by a senior interviewer located in the same call centre. Units selected from the area frame were interviewed by decentralised field interviewers using CAPI. While in some situations field interviewers were permitted to complete some or part of an interview by telephone, roughly three quarters of these interviews were conducted exclusively in person. CAPI interviewers worked independently from their homes using laptop computers and were supervised from a distance by senior interviewers. The variable SAM_TYP on the microdata files indicates whether a case was selected from the area frame (CAPI) or from the telephone or RDD frame (CATI).

In all selected dwellings, a knowledgeable household member was asked to supply basic demographic information on all residents of the dwelling. One member of the household was then selected for a more in-depth interview, which is referred to as the health content interview.

CAPI interviewers were trained to make an initial personal contact with each sampled dwelling. In cases where this initial visit resulted in non-response, telephone follow-ups were permitted. The variable ADM_N09 on the microdata files indicates whether the interview was completed face-to-face, by telephone or using a combination of the two techniques.

To ensure the quality of the data collected, interviewers were instructed to make every effort to conduct the interview with the selected respondent in privacy. In situations where this was unavoidable, the respondent was interviewed with another person present. Flags on the microdata files indicate whether somebody other than the respondent was present during the interview (ADM_N10) and whether the interviewer felt that the respondent's answers were influenced by the presence of the other person (ADM_N11).

To ensure the best possible response rate attainable, many practices were used to minimise non-response, including:

a) Introductory letters

Before the start of each collection period, introductory letters explaining the purpose of the survey were sent to the sampled households. They also explained the importance of the survey and provided examples of how CCHS data would be used.

b) Initiating contact

Interviewers were instructed to make all reasonable attempts to obtain interviews. If the timing of the interviewer's call (or visit) was inconvenient, an appointment was made to call back at a more convenient time. In the case of CAPI interviewers, if requests for appointments were unsuccessful over the telephone, interviewers were instructed to follow-up with a personal visit. If no one was home on the first visit, a brochure with information about the survey and intention to make contact was left at the door. Numerous call-backs were made at different times on different days.

c) Refusal conversion

If an individual initially refused to participate in the survey, a letter was sent from the nearest Statistics Canada Regional Office, stressing the importance of the survey and the household's cooperation. This was followed by a second call (or visit) from a senior interviewer, a project supervisor or another interviewer to try to convince the respondent of the importance of participating in the survey.

d) Language barriers

To remove language as a barrier to conducting interviews, each of the Statistics Canada Regional Offices recruited interviewers with a wide range of language competencies. When necessary, cases were transferred to an interviewer with the language competency needed to complete an interview.

e) Youth interviews

Beginning in 2010, interviewers needed to obtain verbal permission from parents/guardians to interview youths between the ages of 12 to 15 who were selected for interviews. This information was collected in the Parental/Guardian Consent (PGC) block. Several procedures were followed by interviewers to alleviate potential parental concerns and to ensure a completed interview. Interviewers also carried with them a card entitled “Note to parents / guardians about interviewing youths for the Canadian Community Health Survey”. This card explains the purpose of collecting information from youths, lists the subjects to be covered in the survey, encourages giving permission to share and link the obtained information and explains the need to respect a child's right to privacy and confidentiality.

If a parent/guardian asked to see the actual questions, interviewers were instructed to either show the survey questions or, if the interviewer was being conducted by phone, to immediately have the regional office send a copy of the questionnaire.

If privacy could not be obtained to interview the selected youth either in person or over the phone (another person listening in), the interview was coded a refusal. In the case of CAPI interviews, if privacy could not be obtained to interview the selected youth, the interviewer was able to propose to the parent/guardian that the interviewer read the questions out loud and the youth enter their answers directly on the computer.

The Person Most Knowledgeable (PMK) block collected household level information found at the end of the survey (Home Safety, Insurance coverage, Food security, Neurology conditions, Education, Income and Administration) from the most knowledgeable person in the household. This block is initiated when the selected respondent is between the ages of 12 to 17. The block formalizes the process of identifying a person in the household who is likely better able to answer these household level questions than the young selected respondent. If a PMK is found, then the interview moves from the younger selected respondent between the ages of 12 and 17, to a household member who finishes the rest of the interview after the PMK block.

f) Proxy interviews

In cases where the selected respondent was, for reasons of physical or mental health, incapable of completing an interview, another knowledgeable member of the household supplied information about the selected respondent. This is known as a proxy interview. While proxy interviewees were able to provide accurate answers to most of the survey questions, the more sensitive or personal questions were beyond the scope of knowledge of a proxy respondent. This resulted in some questions from the proxy interview being unanswered. Every effort was taken to keep proxy interviews to a minimum.

In 2010, the Proxy interview (GR) block was modified to prompt the interviewer to specifically identify whether the proxy interview was being conducted due to a physical or mental condition. Interviewers were then asked to record the specific condition for either case. The variable ADM_PRX indicates whether a case was completed by proxy.

6.5 Field operations

The majority of the 2012 sample was divided into six two-month collection periods. Regional collection offices were instructed to use the first 4 weeks of each collection period to resolve the majority of the sample, with the next 4 weeks being used to finalise the remaining sample and to follow up on outstanding non-response cases. All cases were to have been attempted by the second week of each collection period.

Sample files were sent approximately two weeks before the start of each collection period to centralised collection offices. A series of dummy cases were included with each CAPI sample. These cases were completed by senior interviewers for the purpose of ensuring that all data

transmission procedures were working through the collection cycle. Once the samples were received, project supervisors were responsible for planning CAPI interviewer assignments. Wherever possible, assignments were generally no larger than 15 cases per interviewer.

Nightly transmission of completed cases from each of the CATI offices to head office was the responsibility of the regional office project supervisor, senior interviewer and the technical support team. Completed CAPI interviews were transmitted daily from the interviewer's home directly to Statistics Canada's head office using a secure telephone transmission.

For final response rates, refer to Appendix E.

6.6 Quality control and collection management

During the collection year, several methods were used to ensure data quality and to optimize collection. These included using internal measures to verify interviewer performance and the use of a series of ongoing reports to monitor various collection targets and data quality.

A system of validation was used for CAPI cases whereby interviewers had their work validated on a regular basis by the regional office. Each collection period, randomly selected cases were flagged in the sample. Regional office managers and supervisors created lists of cases to be validated. These cases were handed to the validation team who then contacted households to verify that a legitimate interview took place. Validation procedures generally occurred during the first few weeks of a collection period to ensure that any issues were detected promptly. Interviewers were provided feedback by their supervisors on a regular basis.

In 2011, an additional quality control system was introduced for CAPI interviews. Upon obtaining consent from the respondents, specific sections of the CAPI interviews were recorded. These recordings were transmitted back to the regional offices, and then randomly chosen for analysis.

CATI interviewers were also randomly chosen for validation. Validation in the CATI collection offices consisted of senior interviewers monitoring interviews to ensure proper techniques and procedures (reading the questions as worded in the applications, not prompting respondents for answers, etc.) were followed by the interviewer.

A series of reports were produced to effectively track and manage collection targets and to assist in identifying other collection issues.

At the end of each collection period, cumulative reports were generated for both the CATI and CAPI samples showing response, link, share and proxy rates by individual health region. The reports were useful in identifying health regions that were below collection target levels, allowing the regional offices to focus efforts in these regions.

Using information obtained from the CAI applications, further analysis was done in head office in order to identify interviews that were completed in less than acceptable time frames. These "short" interviews were flagged, removed from the microdata and treated as non-response.

7. DATA PROCESSING

7.1 Editing

Most editing of the data was performed at the time of the interview by the computer-assisted interviewing (CAI) application. It was not possible for interviewers to enter out-of-range values and flow errors were controlled through programmed skip patterns. For example, CAI ensured that questions that did not apply to the respondent were not asked.

In response to some types of inconsistent or unusual reporting, warning messages were invoked but no corrective action was taken at the time of the interview. Where appropriate, edits were instead developed to be performed after data collection at Head Office. Inconsistencies were usually corrected by setting one or both of the variables in question to "not stated".

7.2 Coding

Pre-coded answer categories were supplied for all suitable variables. Interviewers were trained to assign the respondent's answers to the appropriate category.

In the event that a respondent's answer could not be easily assigned to an existing category, several questions also allowed the interviewer to enter a long-answer text in the "Other-specify" category. All such questions were closely examined in head office processing. For some of these questions, write-in responses were coded into one of the existing listed categories if the write-in information duplicated a listed category. For all questions, the 'Other-specify' responses are taken into account when refining the answer categories for future cycles.

7.3 Creation of derived variables

To facilitate data analysis and to minimize the risk of error, a number of variables on the file have been derived using items found on the CCHS questionnaire. Derived variables generally have a "D", "G" or "F" in the fourth character of the variable name. In some cases, the derived variables are straightforward, involving collapsing of response categories. In other cases, several variables have been combined to create a new variable. The *Derived Variables Documentation (DV)* provides details on how these more complex variables were derived. For more information on the naming convention, please go to Section 12.6.

7.4 Weighting

The principle behind estimation in a probability sample such as CCHS is that each person in the sample "represents", besides himself or herself, several other persons not in the sample. For example, in a simple random 2% sample of the population, each person in the sample represents 50 persons in the population. In the terminology used here, it can be said that each person has a weight of 50.

The weighting phase is a step that calculates, for each person, his or her associated sampling weight. This weight appears on the PUMF, and must be used to derive meaningful estimates from the

survey. For example, if the number of individuals who smoke daily is to be estimated, it is done by selecting the records referring to those individuals in the sample having that characteristic and summing the weights entered on those records.

7.5 Income imputation

Beginning with the 2011 reference year, the household income variable will be imputed. Missing values due to either respondent refusal or respondent's lack of knowledge of household income will be completed using statistical techniques. The main variable of interest is INC_3: 'Total household income - best estimate' but all variables that are derived based on income will also be affected. The income variables along with an imputation flag (INCFIMP in 2011; INCFIMP4 in 2012 and 2011-2012) indicating which values were imputed will be provided on the data file. For more information on the imputation process, please refer to the document 'Income Imputation for the Canadian Community Health Survey' available under the 'Documentation' section of the Definitions, data sources and methods page on the Statistics Canada website or contact Client Services (613-951-1746; fax: 613-951-0792; hd-ds@statcan.gc.ca).

8. WEIGHTING

In order for estimates produced from survey data to be representative of the covered population, and not just the sample itself, users must incorporate the survey weights in their calculations. A survey weight is given to each person included in the final sample, that is, the sample of persons having responded to the survey. This weight corresponds to the number of persons in the entire population that are represented by the respondent.

As described in Section 5, the CCHS has recourse to three sampling frames for its sample selection: an area frame acting as the primary frame and two frames made up of telephone numbers used to complement the area frame. Since only minor differences differentiate the two telephone frames in terms of weighting, they are treated together as one and referred to as being part of the telephone frame.

Depending on the need, one or two frames are used for the selection of the sample within a given health region (HR). When two frames are used, the weighting strategy treats both the area and telephone frames independently to come up with separate household-level weights for each of the frames used. These household-level weights are then combined into a single set of household weights through a step called "*integration*". After applying person-level selection weights and some further adjustments, this integrated weight becomes the final person-level weight.

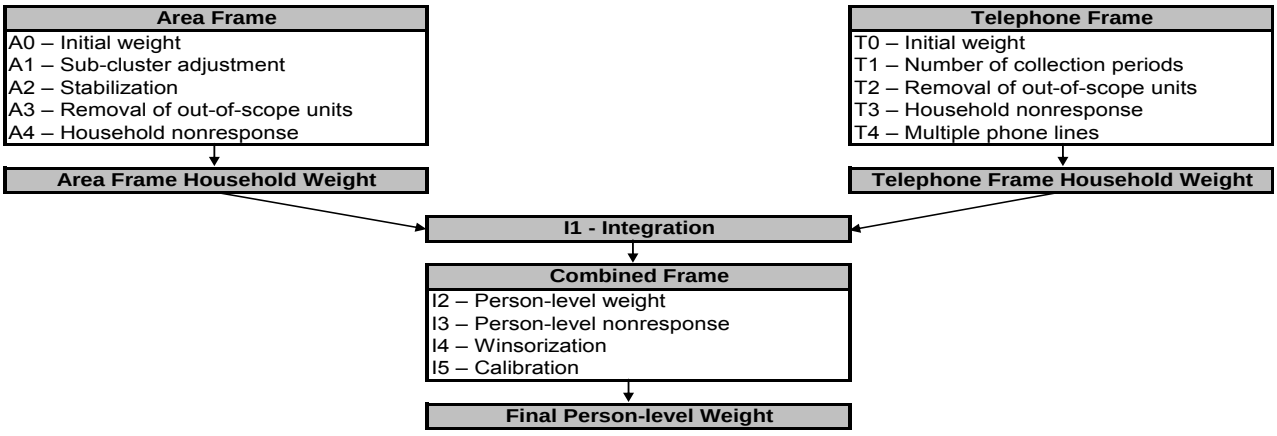
8.1 Overview

As mentioned earlier, units from both the area and telephone frames are treated separately up to the integration step. The following sections describe the weighting process for the provinces. Sub-section 8.2 provides details on the weighting strategy for the area frame, while sub-section 8.3 deals with the strategy for the telephone frame. The integration of the two frames is discussed in 8.4. This is followed by the last weighting steps including calibration, where the weights are adjusted to control for seasonality and to match known population totals. These steps are explained in sub-section 8.5.

Although the two frames are used to cover the three territories, the sampling methods used are slightly different from those used in the provinces. These modifications affect the weighting of these three regions substantially, and they are reported in sub-section 8.6.

Diagram A presents an overview of the different adjustments that are part of the weighting strategy. A numbering system is used to identify each adjustment and will be used throughout the section. Letters *A* and *T* are used as prefixes to refer to adjustments applied to the units on the *A*rea and *T*elephone frames respectively, while prefix *I* identifies adjustments applied from the *I*ntegration step onwards.

Diagram A Weighting strategy overview



8.2 Weighting of the area frame sample

A0 – Initial weight

The weighting on the area frame sample begins with a weight provided by the Labour Force Survey (LFS). This weight is based on the LFS design since the CCHS area frame sample design is based on the LFS. The LFS design consists of a sample of dwellings within clusters selected from LFS strata. In the initial adjustment A0, the LFS weight is adjusted to take into consideration the fact that the CCHS selects a sample to be representative of the Health Region. To do so, the CCHS selects a different number of clusters than the LFS and can repeat the sampling of dwellings within the selected clusters. The resulting weight is called weight A0. For more details about the selection mechanism, as well as a more complete definition of LFS strata and clusters, refer to Statistics Canada (2008)[11].

A1 – Sub-cluster adjustment

In clusters that experience significant growth, a sub-sampling methodology is used to ensure that the workload of the interviewers is kept at a reasonable level. This can consist of sub-sampling from the selected dwellings, dividing the cluster into sub-clusters, or reclassifying the cluster as a stratum and creating new clusters within the stratum. In all these cases, a sub-sample adjustment is calculated and applied to the CCHS weight. This adjustment is applied to weight A0 to produce weight A1. Again, more information can be found in the LFS documentation (Statistics Canada (2008)).

A2 – Stabilization

In some HRs, the increase of the sample size as described in section 5, results in a larger sample than necessary. Stabilization is used to bring the sample size back down to the desired level. The stabilization process consists of randomly sub-sampling dwellings at the HR level from the

[11] Statistics Canada. 2008. *Methodology of the Canadian Labour Force Survey.* Statistics Canada. Cat. No. 71-526-X.

dwellings originally selected within each cluster. An adjustment factor representing the effect of this stabilization is calculated in order to adjust the probability of selection appropriately. This factor, multiplied by weight A1, produces weight A2.

A3 – Removal of out-of-scope units

Among all dwellings sampled, a certain proportion is identified during collection as being out-of-scope. Dwellings that are demolished or under construction, vacant, seasonal or secondary, and institutions are examples of out-of-scope cases for the CCHS. These dwellings and their associated weight are simply removed from the sample. This leaves a sample that consists of, and is representative of, in-scope dwellings or households. These remaining in-scope dwellings maintain the same weight as in the previous step, which is now called weight A3.

A4 – Household nonresponse

During collection, a certain proportion of sampled households inevitably result in nonresponse. This usually occurs when a household refuses to participate in the survey, provides unusable data, or cannot be reached for an interview. Weights of the nonresponding households are redistributed to responding households within response homogeneity groups (RHGs). In order to create the response homogeneity groups, a scoring method based on logistic regression models is used to determine the propensity to respond and these response probabilities are used to divide the sample into groups with similar response properties. The information available for nonrespondents is limited so the regression model uses characteristics such as the collection period and geographic information, as well as paradata or process data, which includes the number of contact attempts, the time/day of attempt, and whether the household was called on a weekend or weekday. Starting in 2008, RHGs were formed within province to better control for provincial totals. An adjustment factor is calculated within each response group as follows:

$$\frac{\textit{Sum of weight A3 for all households}}{\textit{Sum of weight A3 for all responding households}}$$

Weight A3 is multiplied by this factor to produce weight A4 for the responding households. Non-responding households are dropped from the process at this point.

8.3 Weighting of the telephone frame sample

As mentioned earlier, the telephone frame is composed of two frames: a Random Digit Dialling (RDD) frame and a list frame. Only one of the frames can be used within an HR. When the list frame is used, it is always used as a complement to the area frame within the HR. When the RDD frame is used, it is always used as the only frame within the HR. For the purposes of weighting, units coming from the two telephone frames are treated together and therefore are subject to the same adjustments.

The geographical boundaries used to select the sample from the telephone frame do not always conform to the HR geography. Consequently, some units may have been sampled from one HR but

the information collected at the time of the interview places them in a neighbouring HR. This is handled in the weighting by applying the first 3 telephone adjustments (T0, T1 and T2) relative to the HR assigned at the time of sample selection. The remaining 2 adjustments (T3 and T4) are applied to the HR based on information collected from the respondent to ensure that all units belong to their correct HR.

T0 –Initial weight

The initial design weight is defined as the inverse of the probability of selection and is computed separately for the RDD and list frame samples since the method of selection differs between these two frames. For the RDD frame, the selection of telephone numbers is done within each RDD stratum. An RDD stratum is an aggregation of area code prefixes (ACP: the first six digits of a 10-digit telephone number), with each ACP containing valid banks of one hundred numbers (see Norris and Paton[12] for more details). Therefore, the probability of selection is the ratio between the number of sampled units and one hundred times the number of banks within the RDD stratum.

For the list frame, telephone numbers are randomly selected among those assigned to the specific HR. The probability of selection corresponds to the ratio of the number of sampled units to the number of telephone numbers on the list within the HR. The ratio is based on the frame available and the number of units selected for the particular two-month collection period. The probability of selection can therefore change depending on sample allocation and frame updates. The inverse of these probabilities represents the initial weight T0.

T1 – Number of collection periods

On the area frame, the entire sample is selected at the beginning of the year. This is in contrast to the telephone frame, where samples are drawn every two months. Each of these samples comes with an initial weight that allows each sample to be representative of the population at the HR level. To ensure that the total sample represents the population only once, an adjustment factor is applied to reduce the weights of each two-month sample. The adjustment factor applied to each two-month sample is equal to the inverse of the number of samples being combined (i.e. the number of collection periods). Following this adjustment, the entire list frame sample corresponds to the average over the entire combined collection period. The initial weights are multiplied by this adjustment factor to produce weight T1.

T2 - Removal of out-of-scope numbers

Telephone numbers associated with businesses, institutions or other out-of-scope dwellings, as well as numbers not in service or any other non-working numbers are all examples of out-of-scope cases for the telephone frame. Similar to the methods used on the area frame, these cases are simply removed from the process, leaving only in-scope dwellings in the sample. These in-scope dwellings keep the same weight as in the previous step, now called weight T2.

[12] Norris, D.A. and Paton, D.G. 1991. Canada's General Social Survey: Five Years of Experience. *Survey Methodology*. 17, 227-240.

T3 – Household nonresponse

The adjustment applied here to compensate for the effect of household nonresponse is identical to the one applied for the area frame (adjustment A4) although the paradata used does differ because of the differences in collection applications for personal and telephone interviews. The adjustment factor calculated within each response homogeneity group is obtained as follows:

$$\frac{\textit{Sum of weights T2 for all households}}{\textit{Sum of weights T2 for all responding households}}$$

The weight T2 of responding households is multiplied by this factor to produce the weight T3. Nonresponding households are removed from the process at this point.

T4 - Multiple phone lines

Some households can possess more than one residential telephone line. This has an impact on the weighting because these households have a higher probability of being selected. The weights for these households need to be adjusted for the number of residential telephone lines within the household. The adjustment factor represents the inverse of the number of lines in the household. The weight T4 is obtained by multiplying this factor by the weight T3.

8.4 Integration of the telephone and area frames (I1)

This step consists of integrating the weights for households common to the area and telephone frames into a single weight by applying a method of integration[13]. Those units on the area frame that are not on the telephone frame do not have their weights adjusted. For all others units, an adjustment factor α between 0 and 1 is applied to the weights. The weight of the area frame units is multiplied by this factor α, while the weight of the telephone frame units is multiplied by 1- α. Note that in the case where an HR is covered by only one frame, the adjustment factor is equal to 1. Starting in 2008, a fixed α of 0.4 has been used for those units on both frames to ensure greater comparability of estimates across years. The product between the factor derived here and the final household weight calculated earlier (A4 or T4, depending on which frame the unit belongs to), gives the integrated household weight I1.

8.5 Post-integration weighting steps

I2 – Creation of person level weight

Since persons are the desired sampling units, the household-level weights computed to this point need to be converted to the person level. This weight is obtained by multiplying the weight I1 by the

[13] Skinner, C.J. and Rao, J.N.K. 1996. Estimation in Dual Frame Surveys with Complex Designs. *Journal of the American Statistical Association*. 91, 433, 349-356.

inverse of the probability of selection of the person selected in the household. This gives the weight I2. As mentioned earlier, the probability of selection for an individual changes depending on the number of people in the household and the ages of those individuals (see Section 5.6 for more details).

I3 – Person nonresponse

A CCHS interview can be seen as a two-part process. First, the interviewer gets the complete roster of the people within the household. Second, the selected person is interviewed. In some cases, interviewers can only get through the first part, either because they cannot get in touch with the selected person, or because that selected person refuses to be interviewed. Such individuals are defined as person nonrespondents and an adjustment factor must be applied to the weights of person respondents to account for this nonresponse. Using the same methodology that is used in the treatment of household nonresponse, the adjustment is applied within response homogeneity groups. In this process, the scoring method is used to define a response probability based on characteristics available for both respondents and non-respondents. All characteristics collected when creating the roster of household members are available for the estimation of the response probabilities as well as geographic information and some paradata. The probabilities are grouped into response homogeneity groups and the following adjustment factor is calculated within each group:

$$\frac{\textit{Sum of weight I2 for all selected persons}}{\textit{Sum of weight I2 for all responding selected persons}}$$

Weight I2 for responding persons is multiplied by the above adjustment factor to produce weight I3. Nonresponding persons are dropped from the weighting process from this point onward.

I4 – Winsorization

Following the series of adjustments applied to the respondents, some units may come out with extreme weights compared to other units of the same domain of interest. These units could have a large impact on the variance. In order to prevent this, the weight of these outlier units is adjusted downward using a "winsorization" trimming approach.

I5 – Calibration

The last step necessary to obtain the final CCHS weight is calibration (I5). Calibration is done using CALMAR[14] to ensure that the sum of the final weights corresponds to the population estimates defined at the HR level, for all 10 age-sex groups of interest. The five age groups are 12-19, 20-29, 30-44, 45-64, 65+, for both males and females. Starting in 2009, additional controls at more detailed geographic levels were introduced for HRs where additional information is available. A minimum domain size of 20 respondents is required to calibrate at the HR by age by sex level. For domains that have less 20 respondents, some collapsing is done within province and / or within gender. At

[14] Sautory O. CALMAR 2: A New Version of the CALMAR Calibration Adjustment Program. *Proceedings of Statistics Canada Symposium* (Statistics Canada, Catalogue no. 11-522-XCB), 2003.

the same time, weights are adjusted to ensure that each collection period (two-month period) is equally represented within the sample. Note that the calibration is done using the most up to date geography and may not match the geography used in sampling.

The population estimates are based on the 2006 Census counts and counts of birth, death, immigration and emigration since that time. The average of these monthly estimates for each of the HR-age-sex post-strata by collection period is used to calibrate. The weight I4 is adjusted using CALMAR to obtain the final weight I5. Weight I5 corresponds to the ***final CCHS person-level weight*** and can be found on the data file with the variable name WTS_M for master or PUMF users. Prior to the 2010 and 2009-2010 reference period, 2001 Census population counts were used. Evaluation studies have confirmed that the impact of this change on CCHS estimates should be minimal.

8.6 Particular aspects of the weighting in the three territories

As described in Section 5, the sampling frame used in the three territories is somewhat different from the one used in the provinces. Therefore, the weighting strategy is adapted to comply with these differences. This section summarises the changes applied to the steps described in sub-sections 8.1 to 8.5.

For the area frame, as mentioned in sub-section 5.4.1, an additional stage of selection is added in the territories where each territory is stratified into groupings of communities and one community is selected within each group. The capital of each territory forms a stratum on its own and is selected automatically at the first stage. This has an effect in the computation of the probability of selection, and therefore in the value of the initial weight (A0). Once the initial weight is calculated, the same series of adjustments (A1 to A4) is applied to the area frame units. Household-level and person-level nonresponse adjustment classes are built in the same way as for the provinces, using the same set of variables.

For the weighting of the telephone frame units, it should be noted that only the RDD frame is used and its use is exclusive to the capitals of the Yukon and the Northwest Territories. All of the telephone frame adjustments are applied to derive a final weight for the telephone units.
The two sets of weights (area and telephone) are subsequently integrated and calibration is applied in a similar way to what is done for the provinces, with three exceptions. First, the integration is applied only to units located in the Yukon and Northwest Territories capitals since the other communities are covered only by the area frame. Second, the population counts used for calibration for Nunavut represent 70% of the entire population because of the under-coverage of the area frame that was described in section 5.4.1.
Finally, starting with the 2008 and 2007-2008 reference year products, controls have been put in place to ensure that the proportion of aboriginals and the proportion of individuals in the capital regions are controlled in the Northwest Territories and Yukon. A similar control based on Inuit status was introduced for Nunavut. Starting in 2009, the proportion of individuals in the capital regions is controlled in Nunavut. These controls ensure that the proportion of the estimates represented by these different groups is consistent with proportions indicated by the 2006 Census.

8.7 Creation of a share weight

Along with the master file and PUMF which contain all CCHS respondents, a share file is created which contains only a portion (>90%) of the original CCHS respondents. The individuals on this share file have agreed to share their data with certain partners. To compensate for the loss of some respondents from the file, the weights of these "sharers" must be adjusted by the factor:

$$\frac{\text{Sum of weight I3 for all respondents}}{\text{Sum of weight I3 for all respondents agreeing to share their data}}$$

Similar to the nonresponse adjustments, this factor is calculated within homogeneity groups, where in this case, individuals with similar estimated propensity to share will be grouped together. Weight I3 for sharers is multiplied by the above adjustment factor to produce a share weight. Winsorization and calibration, similar to adjustments I4 and I5, are applied to the share units. The final weight after these adjustments is called WTS_S.

8.8 Weighting for a two-year file

When two years of data are combined to create a two-year file, new weights are calculated straightforwardly by halving the annual weights. This ensures that the sum of the final weights is equal to the average population size over the two years. For more information on combining multiple years, please refer to the article "Combining cycles of the Canadian Community Health Survey" published in the Statistics Canada Health Reports publication (82-003) at the following link: http://www.statcan.gc.ca/pub/82-003-x/2009001/article/10795-eng.pdf

9. DATA QUALITY

9.1 Response rates for 2012

In total, 92,682 of the selected units in the CCHS 2012 were in-scope for the survey[15]. Out of these, 71,614 households accepted to participate in the survey resulting in an overall household-level response rate of 77.3%. One individual was selected from each of these 71,614 responding households, out of which a response was obtained for 62,103 individuals, resulting in an overall person-level response rate of 86.7%. At the Canada level, this yields a combined response rate of 67.0% for the CCHS 2012. Table 9.1 provides combined response rates as well as relevant information for their calculation by health region or group of health regions. Table 9.2 provides the same data by Local Health Integrated Network (LHIN).

Table 9.1: 2012 response rates by health region and frames

(see Appendix E)

Table 9.2: 2012 response rates by Local Health Integrated Network (LHIN) and frames in Ontario

(see Appendix E)

9.2 Response rates for 2011-2012

In total, 183,721 of the selected units in the CCHS 2011-2012 were in-scope for the survey. Out of these, 144,000 households accepted to participate in the survey resulting in an overall household-level response rate of 78.4%. One individual was selected from each of these 144,000 responding households, out of which a response was obtained for 125,645 individuals, resulting in an overall person-level response rate of 87.3%. At the Canada level, this yields a combined response rate of 68.4% for the CCHS 2011-2012. Table 9.3 provides combined response rates as well as relevant information for their calculation by health region or group of health regions. Table 9.4 provides the same data by Local Health Integrated Network (LHIN).

Table 9.3: 2011-2012 response rates by health region and frame

(see Appendix G)

[15] Among the units selected, some are not in-scope for the survey. They are, for examples, vacant, demolished or non-residential dwellings or invalid phone numbers such as phone numbers without service or non-residential lines. These units are identified during the data collection, otherwise they would have been excluded before the sample selection. These units are not considered in the calculation of response rates.

Table 9.4: 2011-2012 response rates by Local Health Integrated Network (LHIN) and frame in Ontario

(see Appendix G)

Next, we describe how the various components of the equation should be handled to correctly compute combined response rates.

Household-level response rate

HHRR =	*Number of responding households in both frames* / *All in-scope households in both frames*

Person-level response rate

PPRR =	*Number of responding persons in both frames* / *All selected persons in both frames*

Combined response rate = HHRR x PPRR

Below is an example on how to calculate the combined response rate for Canada using the information found in Table 9.1. The same method applies to rates computed for smaller regions such as province or health region.

HHRR =	$\frac{29{,}396 + 42{,}218}{36{,}175 + 56{,}507}$	=	$\frac{71{,}614}{92{,}682}$	= 0.773
PPRR =	$\frac{26{,}154 + 35{,}949}{29{,}396 + 42{,}218}$	=	$\frac{62{,}103}{71{,}614}$	= 0.867

Combined response rate = 0.773 x 0.867

= 0.6708

= **67.0%**

9.3 Survey Errors

The estimates derived from this survey are based on a sample of individuals. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. than those actually used. The difference between the estimates obtained from the sample and the results from a complete count under similar conditions is called the *sampling error* of the estimate.

Errors which are not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the computer and errors may be introduced in the processing and tabulation of the data. These are all examples of *non-sampling errors*.

9.3.1 Non-sampling Errors

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort was made to reduce non-sampling errors in the CCHS 2010. Quality assurance measures were implemented at each step of data collection and processing to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training with respect to the survey procedures and questionnaire, and the observation of interviewers to detect problems. Testing of the CAI application and field tests were also essential procedures to ensure that data collection errors were minimized.

A major source of non-sampling errors in surveys is the effect of *non-response* on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response. Partial non-response to the CCHS was minimal; once the questionnaire was started, it tended to be completed with very little non-response. Total non-response occurred either because a person refused to participate in the survey or because the interviewer was unable to contact the selected person. Total non-response was handled by adjusting the weight of persons who responded to the survey to compensate for those who did not respond. See section 8 for details on the weight adjustment for non-response.

9.3.2 Sampling Errors

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. The basis for measuring the potential size of sampling errors is the standard deviation of the estimates derived from survey results. However, because of the large variety of estimates that can be produced from a survey, the standard deviation of an estimate is usually expressed relative to the estimate to which it pertains. This resulting measure, known as the coefficient of variation (CV) of an estimate, is obtained by dividing the standard deviation of the estimate by the estimate itself and is expressed as a percentage of the estimate.

For example, suppose hypothetically that it is estimated that 25% of Canadians aged 12 and over are regular smokers and that this estimate is found to have a standard deviation of 0.003. Then the CV of the estimate is calculated as:

$$(0.003/0.25) \times 100\% = 1.20\%$$

Statistics Canada commonly uses CV results when analyzing data and urges users producing estimates from the CCHS data files to also do so. For details on how to determine CVs, see Section 11. For guidelines on how to interpret CV results, see the table at the end of sub-section 10.4.

10. GUIDELINES FOR TABULATION, ANALYSIS AND RELEASE

This section of the documentation outlines the guidelines to be used by users in tabulating, analyzing, publishing or otherwise releasing any data derived from the survey files. With the aid of these guidelines, users of microdata should be able to produce figures that are in close agreement with those produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

10.1 Rounding guidelines

In order that estimates for publication or other release derived from the data files (Master, Share or PUMF) correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates:

a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the proceeding digit is incremented by 1;

b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding;

c) Averages, proportions, rates and percentages are to be computed from unrounded components (i.e., numerators and/or denominators) and then are to be rounded to one decimal using normal rounding. In normal rounding to a single digit, if the final or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is increased by 1;

d) Sums and differences of aggregates (or ratios) are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding;

e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released that differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s);

f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

10.2 Sample weighting guidelines for tabulation

The sample design used for this survey was not self-weighting. That is to say, the sampling weights are not identical for all individuals in the sample. When producing simple estimates, including the production of ordinary statistical tables, users must apply the proper sampling weight. If proper weights are not used, the estimates derived from the data file cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada.

Users should also note that some software packages might not allow the generation of estimates that exactly match those available from Statistics Canada, because of their treatment of the weight field.

10.2.1 Definitions: categorical estimates, quantitative estimates

Before discussing how the survey data can be tabulated and analyzed, it is useful to describe the two main types of point estimates of population characteristics that can be generated from the data files.

Categorical estimates:

Categorical estimates are estimates of the number or percentage of the surveyed population possessing certain characteristics or falling into some defined category. The number of individuals who smoke daily is an example of such an estimate. An estimate of the number of persons possessing a certain characteristic may also be referred to as an estimate of an aggregate.

Example of categorical question:

*At the present do/does ... smoke cigarettes daily, occasionally or not at all? **(SMK_202)***
__ Daily
__ Occasionally
__ Not at all

Quantitative estimates:

Quantitative estimates are estimates of totals or of means, medians and other measures of central tendency of quantities based upon some or all of the members of the surveyed population.

An example of a quantitative estimate is the average number of cigarettes smoked per day by individuals who smoke daily. The numerator is an estimate of the total number of cigarettes smoked per day by individuals who smoke daily, and its denominator is an estimate of the number of individuals who smoke daily.

Example of quantitative question:

*How many cigarettes do/does you/he/she smoke each day now? **(SMK_204)***
|_|_| Number of cigarettes

10.2.2 Tabulation of categorical estimates

Estimates of the number of people with a certain characteristic can be obtained from the data file by summing the final weights of all records possessing the characteristic of interest.

Proportions and ratios of the form $\hat{X} / \hat{Y}$ are obtained by:

a) summing the final weights of records having the characteristic of interest for the numerator ($\hat{X}$);

b) summing the final weights of records having the characteristic of interest for the denominator ($\hat{Y}$); then

c) dividing the numerator estimate by the denominator estimate.

10.2.3 Tabulation of quantitative estimates

Estimates of sums or averages for quantitative variables can be obtained using the following three steps (only step a) is necessary to obtain the estimate of a sum):

a) multiplying the value of the variable of interest by the final weight and summing this quantity over all records of interest to obtain the numerator ($\hat{X}$);

b) summing the final weights of records having the characteristic of interest for the denominator ($\hat{Y}$); then

c) dividing the numerator estimate by the denominator estimate.

For example, to obtain the estimate of the average number of cigarettes smoked each day by individuals who smoke daily, first compute the numerator ($\hat{X}$) by summing the product between the value of variable **SMK_204** and the weight **WTS_M.** Next, sum this value over those records with a value of "daily" to the variable **SMK_202**. The denominator ($\hat{Y}$) is obtained by summing the final weight of those records with a value of "daily" to the variable **SMK_202**. Divide ($\hat{X}$) by ($\hat{Y}$) to obtain the average number of cigarettes smoked each day by daily smokers.

10.3 Guidelines for statistical analysis

The CCHS is based upon a complex design, with stratification and multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used.

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures can differ from what is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variances that are calculated are almost meaningless.
For many analysis techniques (for example linear regression, logistic regression, analysis of variance), a method exists that can make the application of standard packages more meaningful. If the weights on the records are rescaled so that the average weight is one (1), then the results produced by the standard packages will be more reasonable; they still will not take into account the stratification and clustering of the sample's design, but they will take into account the unequal probabilities of selection. The rescaling can be accomplished by using in the analysis a weight equal to the original weight divided by the average of the original weights for the sampled units (people) contributing to the estimator in question.

10.4 Release guidelines

Before releasing and/or publishing any estimate from the data files, users must first determine the number of sampled respondents having the characteristic of interest (for example, the number of respondents who smoke when interested in the proportion of smokers for a given population) in order to ensure that enough observations are available to calculate a quality estimate. For users of the PUMF, if this number is less than 30, the unweighted estimate should not be released regardless of the value of the coefficient of variation for this estimate. For users of the master or share files, it is recommended to have at least 10 observations in the numerator and 20 in the denominator. For weighted estimates, based on sample sizes of 10 or more (30 for the PUMF), users should determine the coefficient of variation of the estimate and follow the guidelines below.

Table 10.1 Sampling variability guidelines

Type of Estimate	CV (in %)	Guidelines
Acceptable	$0.0 \leq CV \leq 16.5$	Estimates can be considered for general unrestricted release. Requires no special notation.
Marginal	$16.6 \leq CV \leq 33.3$	Estimates can be considered for general unrestricted release but should be accompanied by a warning cautioning subsequent users of the high sampling variability associated with the estimates. Such estimates should be identified by the letter E (or in some other similar fashion).
Unacceptable	$CV > 33.3$	Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter F (or in some other fashion) and the following warning should accompany the estimates: "The user is advised that . . .(specify the data) . . . do not meet Statistics Canada's quality standards for this statistical program. Conclusions based on these data will be unreliable and most likely invalid. These data and any consequent findings should not be published. If the user chooses to publish these data or findings, then this disclaimer must be published with the data."

11. APPROXIMATE SAMPLING VARIABILITY TABLES

In order to supply coefficients of variation that will be applicable to a wide variety of categorical estimates produced from a PUMF and that could be readily accessed by the user, a set of Approximate Sampling Variability Tables will be produced with each PUMF. These "look-up" tables allow the user to obtain an approximate coefficient of variation based on the size of the estimate calculated from the survey data.

The coefficients of variation (CV) are derived using the variance formula for simple random sampling and incorporating a factor which reflects the multi-stage, clustered nature of the sample design. This factor, known as the ***design effect,*** was determined by first calculating design effects for a wide range of characteristics and then choosing, for each table produced, a conservative value among all design effects relative to that table. The value chosen was then used to generate a table that applies to the entire set of characteristics.

The Approximate Sampling Variability Tables, along with the design effects, the sample sizes and the population counts that were used to produce them, are provided in the document *Approximate Sampling Variability Tables*, which is available to the share file and PUMF users. All coefficients of variation in the Approximate Sampling Variability Tables are approximate and, therefore, unofficial. Options concerning the computation of exact coefficients of variation are discussed in sub-section 11.7.

Remember: As indicated in Sampling Variability Guidelines in Section 10.4, if the number of observations on which an estimate is based is less than 30, the weighted estimate should not be released regardless of the value of the coefficient of variation. Coefficients of variation based on small sample sizes are too unpredictable to be adequately represented in the tables.

11.1 How to use the CV tables for categorical estimates

The following rules should enable the user to determine the approximate coefficients of variation from the Sampling Variability Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between such estimates.

Rule 1: Estimates of numbers possessing a characteristic (aggregates)

The coefficient of variation depends only on the size of the estimate itself. On the appropriate Approximate Coefficients of Variations Table, locate the estimated number in the left-most column of the table (headed "Numerator of Percentage") and follow the asterisks (if any) across to the first figure encountered. Since not all the possible values for the estimate are available, the smallest value which is the closest must be taken (as an example, if the estimate is equal to 1,700 and the two closest available values are 1,000 and 2,000, the first has to be chosen). This figure is the approximate coefficient of variation.

Rule 2: Estimates of proportions or percentages of people possessing a characteristic

The coefficient of variation of an estimated proportion (or percentage) depends on both the size of the proportion and the size of the numerator upon which the proportion is based. Estimated proportions are relatively more reliable than the corresponding estimates of the numerator of the proportion when the proportion is based upon a sub-group of the population. This is due to the fact that the coefficients of variation of the latter type of estimates are based on the largest entry in a row of a particular table, whereas the coefficients of variation of the former type of estimators are based on some entry (not necessarily the largest) in that same row. (Note that in the tables the CVs decline in value reading across a row from left to right). For example, the estimated proportion of individuals who smoke daily out of those who smoke at all is more reliable than the estimated number who smoke daily.

When the proportion (or percentage) is based upon the total population covered by each specific table, the CV of the proportion is the same as the CV of the numerator of the proportion. In this case, this is equivalent to applying Rule 1.

When the proportion (or percentage) is based upon a subset of the total population (e.g., those who smoke at all), reference should be made to the proportion (across the top of the table) and to the numerator of the proportion (down the left side of the table). Since not all the possible values for the proportion are available, the smallest value which is the closest must be taken (for example, if the proportion is 23% and the two closest values available in the column are 20% and 25%, 20% must be chosen). The intersection of the appropriate row and column gives the coefficient of variation.

Rule 3: Estimates of differences between aggregates or percentages

The standard error of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard error considered separately. That is, the standard error of a difference ($\hat{d} = \hat{X}_2 - \hat{X}_1$) is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

where $\hat{X}_1$ is estimate 1, $\hat{X}_2$ is estimate 2, and α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d}$. This formula is accurate for the difference between independent populations or subgroups, but is only approximate otherwise. It will tend to overstate the error, if $\hat{X}_1$ and $\hat{X}_2$ are positively correlated and understate the error if $\hat{X}_1$ and $\hat{X}_2$ are negatively correlated.

Rule 4: Estimates of ratios

In the case where the numerator is a subset of the denominator, the ratio should be converted to a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is

the number of individuals who smoke at all and the numerator is the number of individuals who smoke daily out of those who smoke at all.

Consider the case where the numerator is not a subset of the denominator, as for example, the ratio of the number of individuals who smoke daily or occasionally as compared to the number of individuals who do not smoke at all. The standard deviation of the ratio of the estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately multiplied by $\hat{R}$, where $\hat{R}$ is the ratio of the estimates ($\hat{R} = \hat{X}_1 / \hat{X}_2$). That is, the standard error of a ratio is:

$$\sigma_{\hat{R}} = \hat{R}\sqrt{\alpha_1^{\,2} + \alpha_2^{\,2}}$$

Where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively.

The coefficient of variation of $\hat{R}$ is given by $\sigma_{\hat{R}} / \hat{R} = \sqrt{\alpha_1^{\,2} + \alpha_2^{\,2}}$. The formula will tend to overstate the error, if $\hat{X}_1$ and $\hat{X}_2$ are positively correlated and understate the error if $\hat{X}_1$ and $\hat{X}_2$ are negatively correlated.

Rule 5: Estimates of differences of ratios

In this case, Rules 3 and 4 are combined. The CVs for the two ratios are first determined using Rule 4, and then the CV of their difference is found using Rule 3.

11.2 Examples of using the CV tables for categorical estimates

The following "real life" examples are included to assist users in applying the foregoing rules.

Example 1: Estimates of numbers possessing a characteristic (aggregates)

Suppose that a user estimates that 4,722,617 individuals smoke daily in Canada. How does the user determine the coefficient of variation of this estimate?

1) Refer to the CANADA level CV table.

2) The estimated aggregate (4,722,617) does not appear in the left-hand column (the "Numerator of Percentage" column), so it is necessary to use the smallest figure closest to it, namely 4,000,000.

3) The coefficient of variation for an estimated aggregate (expressed as a percentage) is found by referring to the first non-asterisk entry on that row, namely, 1.7%.

4) So the approximate coefficient of variation of the estimate is 1.7%. According to the Sampling Variability Guidelines presented in Section 10.4, the finding that there were 4,722,617 individuals who smoke daily is publishable with no qualifications.

Example 2 : Estimates of proportions or percentages possessing a characteristic

Suppose that the user estimates that 4,722,617/6,081,453=77.7% of individuals in Canada who smoke at all smoke daily. How does the user determine the coefficient of variation of this estimate?

1) Refer to the CANADA level CV table.

2) Because the estimate is a percentage which is based on a subset of the total population (i.e., individuals who smoke at all, that is to say, daily or occasionally), it is necessary to use both the percentage (77.7%) and the numerator portion of the percentage (4,722,617) in determining the coefficient of variation.

3) The numerator (4,722,617) does not appear in the left-hand column (the "Numerator of Percentage" column) so it is necessary to use the smallest figure closest to it, namely 4,000,000. Similarly, the percentage estimate does not appear as any of the column headings, so it is necessary to use the figure closest to it, 70.0%.

4) The figure at the intersection of the row and column used, namely 1.0% is the coefficient of variation (expressed as a percentage) to be used.

5) So the approximate coefficient of variation of the estimate is 1.0%. According to the Sampling Variability Guidelines presented in Section 10.4, the finding that 77.7% of individuals who smoke at all smoke daily can be published with no qualifications.

Example 3 : Estimates of differences between aggregates or percentages

Suppose that a user estimates that, among men, 2,535,367/13,078,499 = 19.4% smoke daily (estimate 1), while for women, this percentage is estimated at 2,187,250 / 13,476,931 = 16.2% (estimate 2). How does the user determine the coefficient of variation of the difference between these two estimates?

1) Using the CANADA level CV table in the same manner as described in example 2 gives the CV for estimate 1 as 2.4% (expressed as a percentage), and the CV for estimate 2 as 2.4% (expressed as a percentage).

2) Using rule 3, the standard error of a difference ($\hat{d} = \hat{X}_2 - \hat{X}_1$) is :

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

Where $\hat{X}_1$ is estimate 1, $\hat{X}_2$ is estimate 2, and α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The standard error of the difference $\hat{d}$ = (0.194 - 0.162) = 0.032 is :

$$\sigma_{\hat{d}} = \sqrt{[(0.194)(0.024)]^2 + [(0.162)(0.024)]^2}$$
$$= 0.0061$$

3) The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d}$ = 0.0061/0.032 = 0.190.

4) So the approximate coefficient of variation of the difference between the estimates is 19.0% (expressed as a percentage). According to the Sampling Variability Guidelines presented in Section 10.4, this estimate can be published but a warning has to be issued.

Example 4 : Estimates of ratios

Suppose that the user estimates that 4,722,617 individuals smoke daily, while 1,358,836 individuals smoke occasionally. The user is interested in comparing the estimate of daily to occasional smokers in the form of a ratio. How does the user determine the coefficient of variation of this estimate?

1) First of all, this estimate is a ratio estimate, where the numerator of the estimate (= $\hat{X}_1$) is the number of individuals who smoke occasionally. The denominator of the estimate (= $\hat{X}_2$) is the number of individuals who smoke daily.

2) Refer to the CANADA level CV table.

3) The numerator of this ratio estimate is 1,358,836. The smallest figure closest to it is 1,000,000. The coefficient of variation for this estimate (expressed as a percentage) is found by referring to the first non-asterisk entry on that row, namely, 3.7%.

4) The denominator of this ratio estimate is 4,722,617. The figure closest to it is 4,000,000. The coefficient of variation for this estimate (expressed as a percentage) is found by referring to the first non-asterisk entry on that row, namely, 1. 7%.

5) So the approximate coefficient of variation of the ratio estimate is given by rule 4, which is,

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^{\,2} + \alpha_2^{\,2}},$$

That is,

$$\alpha_{\hat{R}} = \sqrt{(.037)^2 + (.017)^2}$$

$$= 0.041$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The obtained ratio of occasional to daily smokers is 1,358,836/4,722,617 which is 0.29:1. The coefficient of variation of this estimate is 4.1% (expressed as a percentage), which is releasable with no qualifications, according to the Sampling Variability Guidelines presented in Section 10.4.

11.3 How to use the CV tables to obtain confidence limits

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example a 95% confidence interval can be described as follows: if sampling of the population is repeated indefinitely, each sample leading to a new confidence interval for an estimate, then in 95% of the samples the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about 99 out of 100 that the differences would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate, $\hat{X}$, are generally expressed as two numbers, one below the estimate and one above the estimate, as $\left(\hat{X}-k,\ \hat{X}+k\right)$, where k is determined depending upon the level of confidence desired and the sampling error of the estimate.

Confidence intervals for an estimate can be calculated directly from the Approximate Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate $\hat{X}$, and then using the following formula to convert to a confidence interval CI:

$$CI_X = [\ \hat{X} - z\ \hat{X}\ \alpha_{\hat{X}}\ ,\hat{X} + z\ \hat{X}\ \alpha_{\hat{X}}\ ]$$

Where $\alpha_{\hat{X}}$ is determined coefficient of variation for $\hat{X}$, and

$z =$ 1 if a 68% confidence interval is desired
$z =$ 1.6 if a 90% confidence interval is desired
$z =$ 2 if a 95% confidence interval is desired
$z =$ 3 if a 99% confidence interval is desired.

Note: Release guidelines presented in section 10.4 which apply to the estimate also apply to the confidence interval. For example, if the estimate is not releasable, then the confidence interval is not releasable either.

11.4 Example of using the CV tables to obtain confidence limits

A 95% confidence interval for the estimated proportion of individuals who smoke daily from those who smoke at all (from example 2, sub-section 11.2) would be calculated as follows:

$\hat{X}$ = 0.777

z = 2

$\alpha_{\hat{X}}$ = 0.01 is the coefficient of variation of this estimate as determined from the tables.

$CI_{\hat{x}}$ = {0.777 - (2) (0.777) (0.01) , 0.777 + (2) (0.777) (0.01)}

$CI_{\hat{x}}$ = {0.761, 0.793}

11.5 How to use the CV tables to do a Z-test

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let $\hat{X}_1$ and $\hat{X}_2$ be sample estimates for 2 characteristics of interest. Let the standard error on the difference $\hat{X}_1 - \hat{X}_2$ be $\sigma_{\hat{d}}$. If the ratio of $\hat{X}_1 - \hat{X}_2$ over $\sigma_{\hat{d}}$ is between -2 and 2, then no conclusion about the difference between the characteristics is justified at the 5% level of significance. If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 0.05 level.

11.6 Example of using the CV tables to do a Z-test

Let us suppose we wish to test, at 5% level of significance, the hypothesis that there is no difference between the proportion of men who smoke daily AND the proportion of women who smoke daily. From example 3, sub-section 11.2, the standard error of the difference between these two estimates was found to be = 0.00461. Hence,

$$z = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}} = \frac{0.194 - 0.162}{0.0061} = \frac{0.032}{0.0061} = 5.25$$

Since $z = 5.25$ is greater than 2, it must be concluded that there is a significant difference between the two estimates at the 0.05 level of significance. Note that the two sub-groups compared are considered as being independent, so the test is correct.

11.7 Exact variances/coefficients of variation

All coefficients of variation in the Approximate Sampling Variability Tables (CV Tables) are indeed approximate and, therefore, unofficial.

The computation of exact coefficients of variation is not a straightforward task since there is no simple mathematical formula that would account for all CCHS sampling frame and weighting aspects. Therefore, other methods such as resampling methods must be used in order to estimate measures of precision. Among these methods, the bootstrap method is the one recommended for analysis of CCHS data.

The computation of coefficients of variation (or any other measure of precision) with the use of the bootstrap method requires access to information that is considered confidential and not available on the PUMF. This computation must be done using the Master file. Access to the Master file is discussed in section 12.3.

For the computation of coefficients of variation, the bootstrap method is advised. A macro program, called "Bootvar", was developed in order to give users easy access to the bootstrap method. The Bootvar program is available in SAS and SPSS formats, and is made up of macros that calculate the variances of totals, ratios, differences between ratios, and linear and logistic regressions.

There are a number of reasons why a user may require an exact variance. A few are given below.

Firstly, if a user desires estimates at a geographic level other than those available in the tables (for example, at the rural/urban level), then the CV tables provided are not adequate. Coefficients of variation of these estimates may be obtained using "domain" estimation techniques through the exact variance program.

Secondly, should a user require more sophisticated analyses such as estimates of parameters from linear regressions or logistic regressions, the CV tables will not provide correct associated coefficients of variation. Although some standard statistical packages allow sampling weights to be incorporated in the analyses, the variances that are produced often do not take into account the stratified and clustered nature of the design properly, whereas the exact variance program would do so.

Thirdly, for estimates of quantitative variables, separate tables are required to determine their sampling error. Since most of the variables for the CCHS are primarily categorical in nature, this has not been done. Thus, users wishing to obtain coefficients of variation for quantitative variables can do so through the exact variance program. As a general rule, however, the coefficient of variation of a quantitative total will be larger than the coefficient of variation of the corresponding category estimate (i.e., the estimate of the number of persons contributing to the quantitative estimate). If the corresponding category estimate is not releasable, the quantitative estimate will not be either. For example, the coefficient of variation of the estimate of the total number of cigarettes smoked each day by individuals who smoke daily would be greater than the coefficient of variation of the corresponding estimate of the number of individuals who smoke daily. Hence if the

coefficient of variation of the latter is not releasable, then the coefficient of variation of the corresponding quantitative estimate will also not be releasable.

Lastly, should users find themselves in a position where they can use the CV tables, but this renders a coefficient of variation in the "marginal" range (16.6% - 33.3%), the user should release the associated estimate with a warning cautioning users of the high sampling variability associated with the estimate. This would be a good opportunity to recalculate the coefficient of variation through the exact variance program to find out if it is releasable without a qualifying note. The reason for this is that the coefficients of variation produced by the tables are based on a wide range of variables and are therefore considered crude, whereas the exact variance program would give an exact coefficient of variation associated with the variable in question.

11.8 Release cut-offs for the CCHS

The document *Approximate Sampling Variability Table*, which is available to the share file and PUMF users, presents tables giving the minimum cut-offs for estimates of totals at the Canada, provincial, health region and CLSC levels and those for various age groups at the Canada level. Estimates smaller than the value given in the "Marginal" column may not be released under any circumstances.

12. MICRODATA FILES: DESCRIPTION, ACCESS AND USE

The CCHS produces three types of microdata files: master files, share files and public use microdata files (PUMF). Table 12.1 includes the list of all available 2012 and 2011-2012 data files.

12.1 Master files

The master files contain all variables and all records from the survey collected during a collection period. These files are accessible at Statistics Canada for internal use and in Statistics Canada's Research Data Centres (RDC), and are also subject to custom tabulation requests.

12.1.1 Research Data Centre

The RDC Program enables researchers to use the survey data in the master files in a secure environment in several universities across Canada. Researchers must submit research proposals that, once approved, give them access to the RDC. For more information, please consult the following web page: http://www.statcan.ca/english/rdc/index.htm

12.1.2 Custom tabulations

Another way to access the master files is to offer all users the option of having staff in Client Services of the Health Statistics Division prepare custom tabulations. This service is offered on a cost-recovery basis. It allows users who do not possess knowledge of tabulation software products to get custom results. The results are screened for confidentiality and reliability concerns before release. For more information, please contact Client Services at 613-951-1746 or by e-mail at hd-ds@statcan.gc.ca.

12.1.3 Remote access

Finally, the remote access service to the survey master files is another way to have access to these data if, for some reason, the user cannot access a Research Data Centre. Each purchaser of the microdata product can be supplied with a synthetic or 'dummy' master file and a corresponding record layout. With these tools, the researcher can develop his own set of analytical computer programs. The code for the custom tabulations is then sent via e-mail to cchs-escc@statcan.gc.ca. The code will then be transferred into Statistics Canada's internal secured network and processed using the appropriate master file of CCHS data. Estimates generated will be released to the user, subject to meeting the guidelines for analysis and release outlined in Section 10 of this document. Results are screened for confidentiality and reliability concerns and then the output is returned to the client. There is no charge for this service.

12.2 Share files

The share files contain all variables and all records of CCHS respondents who agreed to share their data with Statistic Canada's partners, which are the provincial and territorial health departments, Health Canada and the Public Health Agency of Canada. Statistics Canada also asks respondents living in Quebec for their permission to share their data with the Institut de la statistique du Québec. The share file is released only to these organizations. Personal identifiers are removed from the share files to respect respondent confidentiality. Users of these files must first certify that they will not disclose, at any time, any information that might identify a survey respondent.

12.3 Public use microdata files

The public use microdata files (PUMF) are developed from the master files using a technique that balances the need to ensure respondent confidentiality with the need to produce the most useful data possible at the health region level. The PUMF must meet stringent security and confidentiality standards required by the *Statistics Act* before they are released for public access. To ensure that these standards have been achieved, each PUMF goes through a formal review and approval process by an executive committee of Statistics Canada.

Variables most likely to lead to identification of an individual are deleted from the data file or are collapsed to broader categories.

The PUMF contains the data collected over two years. It includes questions that were asked over two years. Unless otherwise specified, these questions are usually those included in the annual common content and in the two-year common content as well as the optional content selected for two years by the provinces and territories.

There is no charge to access the PUMF in a post-secondary educational institution that is part of the Data Liberation Initiative. They are also free of charge from Client Services on request at 613-951-1746 or by e-mail at hd-ds@statcan.gc.ca.

Table 12.1 2012 and 2011-2012 CCHS data files

Reference period	Files	File name	Sampling weight	Bootstrap weights file	Variables included	Records included
2012	Main master file	HS.txt	WTS_M	b5.txt	All common and all optional modules.	All respondent records
	Share file	HS.txt	WTS_S	b5.txt	All common and all optional modules.	Records of all respondents who agreed to share their data
2011-2012	Main master file	HS.txt	WTS_M	b5.txt	All common and all optional modules.	All respondent records
	Share file	HS.txt	WTS_S	b5.txt	All common and all optional modules.	Records of all respondents who agreed to share their data

12.4 How to use the CCHS data files: annual data file or two-year data file?

Since the 2008 and 2007-2008 data were released, users that have access to share files or master files have had the choice of using one-year or two-year data files. Decisions about which period to use in a given data analysis should be guided by the level of detail and the quality required. With a one-year file, estimates will not always be available because of the quality associated with limited sample sizes.

Before interpreting and using a CCHS estimate, it is recommended to make sure that the estimates meets the following rules:

- Coefficient of Variation 33.3% or less
- a minimum of 10 respondents in the domain with the characteristic and
- total domain of interest includes at least 20 respondents.

This will not be possible for rare characteristics and detailed domains with one-year files. Instead, users will have to rely on two-year files or multi-year files.

Where the use of either a one-year or two-year file is viable, the user should consider the trade-off between accuracy and currency. If it is important to reflect the current characteristics of a population as closely as possible, the one-year file would be preferable. However, with the increased sample size, more detailed estimates and analyses can be carried out with a two-year file.

12.5 Use of weight variable

The weight variable **WTS_M** represents the sampling weight for key survey files. For a given respondent, the sampling weight can be interpreted as the number of people the respondent represents in the Canadian population. This weight must always be used when computing statistical estimates in order to make inference at the population level possible. The production of unweighted estimates is not recommended. The sample allocation, as well as the survey design specifics can cause such results to not correctly represent the population. Refer to section 8 on weighting for a more detailed explanation on the creation of this weight. The weight variable **WTS_M** must be used for regional analyses.

The **Food Security** module, included in certain reference period data files, measures concepts that apply not only to the respondent's situation, but also to that of the respondent's entire household. Depending on the level of analysis, the analysis of the variables may require use of a weight calculated to represent the number of Canadian households, rather than the number of persons. This weight variable **WTS_HH** is found in a separate file (HS_HHWT.txt). It can be used in place of the variable **WTS_M** for household analyses at the national and provincial levels.

12.6 Variable naming convention beginning in 2007

The variable naming convention adopted allows data users to easily use and identify the data based on the module and variable type. The CCHS variable naming convention fulfils two requirements: to restrict variable names to a maximum of eight characters for ease of use by analytical software products and to identify easily conceptually identical variables from one survey collection period to the next. Questions to which changes are made between two collection periods, and where the changes alter the concept measured by the question, are entirely renamed to avoid any confusion in the analysis.

The CCHS variable naming convention was changed beginning with the data from the 2007 collection period. The letter corresponding to the survey version (e.g., A = 2000 (cycle 1.1), C = 2003 (cycle 2.1) and E = 2005 (3.1) is no longer used in the variable names. A new variable (REFPER, format = YYYYMM-YYYYMM) was added to the microdata files in order to identify the beginning and the end of the reference during which data included in the file were collected. This variable will be useful, notably for users wanting to use data from several collection periods at a time. Therefore, variable names for identical modules or questions from one collection year to the next (e.g., 2007 and 2008) will be the same.

The naming convention used for variables beginning with the 2007 CCHS use up to eight characters. The variable names are structured as follows:

Positions 1 to 3:	Module/questionnaire section name
Position 4:	Variable type (underscore, C, D, F or G)
Positions 5 to 8:	Question number and answer option for multiple response questions

Example 1 shows that the structure of the variable name for question 202, Smoking Module, is SMK_202 :

Positions 1 to 3:	SMK	Smoking module
Position 4 :	_	(underscore = collected data)
Position 5 to 8:	202	Question number

Example 2 shows the structure of the variable name for question 2 of the Health Care Utilization Module (HCU_02A), which is a multi-response question:

Positions 1 to 3:	**HCU**	Health care utilization module
Position 4 :	_	(underscore = collected data)
Position 5 to 8:	02AA	Corresponding question number and answer option

Positions 1 to 3 contain the acronyms for each of the modules. These acronyms appear beside the module names given in the table in Appendix A.

Position 4 designates the variable type based on whether it is a variable collected directly from a questionnaire question ("_"), from a coded ("C"), derived ("D"), grouped ("G"), or flag ("F") variable.

In general, the last four positions (5 to 8) follow the variable numbering used on the questionnaire. The letter "Q" used to represent the word "question" is removed, and all question numbers are presented in a two or three digit format. For example, question Q01A in the questionnaire becomes simply 01A, and question Q15 becomes simply 15.

Table 12.2 Designation of codes used in the 4th position of the CCHS variable names

_	Collected variable	A variable that appears directly on the questionnaire
C	Coded variable	A variable coded from one or more collected variables (e.g., SIC, Standard Industrial Classification code)
D	Derived variable	A variable calculated from one or more collected or coded variables, usually calculated during head office processing (e.g., Health Utility Index)
F	Flag variable	A variable calculated from one or more collected variables (like a derived variable), but usually calculated by the data collection computer application for later use during the interview (e.g., work flag)
G	Grouped variable	Collected, coded, suppressed or derived variables collapsed into groups (e.g., age groups)

For questions that have more than one response option, the final position in the variable naming sequence is represented by a letter. For this type of question, new variables were created to differentiate between a "yes" or "no" answer for each response option. For example, if Q2 had 4 response options, the new questions would be named Q2A for option 1, Q2B for option 2, Q2C for

option 3, etc. If only options 2 and 3 were selected, then Q2A = No, Q2B = Yes, Q2C = Yes and Q2D = No.

12.7 Variable naming convention before 2007

As mentioned earlier, the variable naming convention was changed in 2007. The flag for the cycle in which the variables were collected was removed. This flag was found in the 4th position for 2000 to 2005 data (cycles 1.1 to 3.1).

Here is the list of letters used in the CCHS microdata files between cycles 1.1 and 3.1 and their corresponding cycle.

Letter	Cycle and cycle name
A	**2000 (**Cycle 1.1) : Canadian Community Health Survey
B	**2002 (**Cycle 1.2) : Canadian Community Health Survey - Mental Health and Well-Being
C	**2003 (**Cycle 2.1) : Canadian Community Health Survey
D	**2004 (**Cycle 2.2) : Canadian Community Health Survey - Nutrition
E	**2005 (**Cycle 3.1) : Canadian Community Health Survey

12.8 Guidelines for the use of sub-sample variables – Not applicable to 2012 and 2011-2012 data files

12.9 Data dictionaries

Separate data dictionary reports, including universe statements and frequencies, are provided for the main master and share files and each of the sub-sample files.

In the master file data dictionary reports, optional content modules are treated in the same way as previous CCHS cycles. For each module, a flag indicates whether a given respondent lives in a health region where the module was selected as optional content. When the flag is equal to 2 (No), all variables in the module have “not applicable” values. For example, the DOWST variable indicates if the Work stress module applies to a given respondent.

12.10 Differences in calculation of common content variables using different files

Variables from common content modules can be estimated using either of the two data files provided, when a one year and a two-year data file is available. Depending on which file is used, very small differences will be observed.

All official Statistics Canada estimates of variables from common modules are based on the main master file sampling weight.

Appendix A – Canadian community health survey content overview (2011-2012)

Appendix A – Canadian community health survey content overview (2011 - 2012)

Annual common content (all health regions)[i]

- Age of respondent (ANC)
- Alcohol use (ALC)
- Chronic conditions (CCC)
- Contact with health professionals (CHP)[iii]
- Exposure to second-hand smoke (ETS)
- Fruit and vegetable consumption (FVC)
- Flu shots (FLU)
- General health (GEN)
- Health care utilization (HCU)
- Height and weight – Self –reported (HWT)
- Pain and discomfort (HUP)
- Physical activities (PAC)
- Smoking (SMK)

Administration and Socio-demographics

- Administration information (ADM)
- Education (EDU)
- Income (INC)
- Labour force (LBS)
- Person most knowledgeable about the household situation (PMK)
- Socio-demographic characteristics (SDC)

Two-year biennial common content (all health regions)

2011-2012

1) *Healthy Living*

- Sedentary activities (SAC)
- Food security (FSC)
- Maternal experiences - breastfeeding (MEX)

2) Mental well-being

- Consultation about mental health (CMH)
- Positive mental health (PMH)

One-year biennial/quadrennial common content (all health regions)

2011

1) Health Services Access Survey [ii]

- Access to health care services (ACC)
- Waiting times (WTM)

2) Neurological Conditions (NEU)

2012

1) *Chronic Disease Screening*

- PAP smear test (PAP)
- Mammography (MAM)
- Colorectal cancer screening (CCS)
- Spirometry (SPI)
- Blood test (BLT)
- Physical check-up (PCU)

Rapid Response (national estimates only)

2011

- Neighbourhood environment (NBE) (July – August)

2012

- Canada's food guide (CFG) (May – June)
- Food Skills – Part 1 (FS1) (November-December)

i RAC has been a core module throughout the years, with the exception of 2011 when it was asked only in the territories.
ii Asked of a sub-sample of respondents. These theme modules were not asked of respondents in the territories.
iii In 2011, CHP changed from being a common content module to an optional module. In 2012, CHP returned to being a common content module but was divided into two modules (CHP and CP2), CP2 is an optional module.

Appendix B – Selection of optional content by province and territory (2012 and 2011-2012)

Appendix B – Optional content selection by health regions (grouped by province) (2012)

Description	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt
ACC Access to health care services	☑	☐	☐	☑	☐	☐	☐	☐	☐	☑	☐	☐	☐
ADL Activites of Daily Living	☑	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐
ALW Alcohol use during the past week	☑	☐	☐	☐	☑	☑	☑	☑	☐	☐	☐	☐	☑
BPC Blood pressure check	☑	☑	☐	☐	☐	☐	☑	☐	☑	☐	☐	☑	☐
BRX Breast examinations	☐	☑	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑
BSX Breast self-examinations	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☑
CIH Changes made to improve health	☑	☑	☑	☐	☑	☐	☑	☐	☑	☐	☑	☑	☐
DEN Dental visits	☑	☐	☐	☐	☐	☐	☑	☐	☑	☐	☐	☑	☐
DEP Depression	☑	☑	☑	☑	☐	☐	☐	☑	☑	☐	☐	☐	☑
DIA Diabetes care	☑	☐	☑	☑	☐	☑	☐	☐	☐	☐	☐	☐	☐
DSU Dietary supplement use - Vitamins and minerals	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☑
DIS Distress	☐	☐	☐	☐	☐	☑	☐	☑	☐	☐	☐	☐	☐
DRV Driving and Safety	☐	☑	☐	☐	☐	☐	☑	☐	☑	☐	☐	☑	☐
FDC Food choices	☐	☑	☑	☐	☐	☐	☐	☑	☑	☑	☐	☑	☐
HCS Health care system satisfaction	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☑
HUI Health utilities index	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐
HMC Home care services	☑	☐	☐	☑	☐	☑	☐	☐	☑	☐	☐	☐	☐
HMS Home safety	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑
HRT Hormone replacement therapy	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
IDG Illicit drugs use	☐	☐	☐	☐	☐	☑	☐	☑	☐	☐	☐	☐	☑
INJ Injuries	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑	☐

Description	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt
LOP Loss of productivity	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐
MAS Mastery	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐
MXA Maternal experiences – Alcohol use during pregnancy	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑
MXS Maternal experiences – Smoking during pregnancy	☑	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑
MED Medication use	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐
OH1 Oral health 1	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐
OH2 Oral health 2	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐
PSC Patient satisfaction - Community Based Care	☐	☑	☐	☐	☐	☑	☐	☐	☐	☑	☑	☐	☐
PAS Patient satisfaction - Health Care Services	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐
SCP Physical activity - Stages of change	☑	☐	☑	☐	☑	☐	☐	☐	☐	☑	☐	☑	☐
PSA Prostate cancer screening	☐	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☑	☐
SWL Satisfaction with life	☐	☐	☐	☐	☐	☐	☐	☐	☑	☑	☐	☐	☐
SFE Self-esteem	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑
SXB Sexual behaviours	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☑	☑
SLP Sleep	☐	☐	☑	☐	☑	☐	☑	☐	☑	☐	☑	☐	☐
TAL Smoking - Other tobacco products	☐	☐	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☐
SPC Smoking - Physician counselling	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑
SCH Smoking - Stages of change	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑	☑	☐	☐
YSM Smoking - Youth smoking	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑	☐	☐
SCA Smoking cessation methods	☐	☐	☑	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑
SPS Social provisions	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑	☑	☑	☑
SSA Social Support - Availability	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐
SSU Social Support - Utilization	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐

Description	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt
SPR Spiritual values	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☐
CST Stress - Child and adult stressors	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐
RLE Stress - Recent life events	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐
STS Stress - Sources	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☑	☐	☐
SUI Suicidal Thoughts and Attempts	☑	☐	☐	☐	☐	☐	☐	☑	☑	☐	☐	☐	☑
SSB Sun safety behaviours	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐
UCN Unmet health care needs	☐	☐	☑	☐	☐	☑	☐	☐	☐	☑	☐	☑	☑
UPE Use of Protective Equipment	☑	☑	☐	☐	☐	☑	☐	☐	☑	☐	☑	☐	☐
ORG Voluntary organizations - Participation	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☐	☑	☐
WTM Waiting Times	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Appendix B – Optional content selection by health regions (grouped by province) (2011-2012)

Description	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt
ACC Access to health care services	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐
ADL Activites of Daily Living	☑	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐
ALW Alcohol use during the past week	☑	☐	☐	☐	☑	☑	☑	☑	☐	☐	☐	☐	☐
BPC Blood pressure check	☑	☑	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑	☐
BLT Blood test	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☑
BRX Breast examinations	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑
BSX Breast self-examinations	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☑
CIH Changes made to improve health	☑	☑	☑	☐	☑	☐	☑	☐	☑	☐	☑	☑	☐
CCS Colorectal Cancer Screening	☑	☑	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☑
CHP Contacts with health professionals	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐
DEN Dental visits	☑	☐	☐	☐	☐	☐	☑	☐	☑	☐	☐	☑	☐
DEP Depression	☑	☑	☑	☑	☐	☐	☐	☑	☐	☐	☐	☐	☑
DIA Diabetes care	☑	☐	☑	☑	☐	☑	☐	☐	☐	☐	☐	☐	☐
DSU Dietary supplement use - Vitamins and minerals	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☑
DIS Distress	☐	☐	☐	☐	☐	☑	☐	☑	☐	☐	☐	☐	☐
DRV Driving and Safety	☐	☑	☐	☐	☐	☐	☑	☐	☑	☐	☐	☑	☐
FDC Food choices	☐	☑	☑	☐	☐	☐	☐	☑	☑	☑	☐	☑	☐
HCS Health care system satisfaction	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☑
HMC Home care services	☑	☐	☐	☑	☐	☑	☐	☐	☑	☐	☐	☐	☐
HMS Home safety	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑
HRT Hormone replacement therapy	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Description		N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt
IDG	Illicit drugs use	☐	☐	☐	☐	☐	☑	☐	☑	☐	☐	☐	☐	☑
INJ	Injuries	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑	☐
LOP	Loss of productivity	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐
MAM	Mammography	☑	☐	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☑
MAS	Mastery	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐
MXA	Maternal experiences – Alcohol use during pregnancy	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑
MXS	Maternal experiences – Smoking during pregnancy	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑
MED	Medication use	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐
OH1	Oral health 1	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐
OH2	Oral health 2	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐
PAP	PAP Smear Test	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐
PSC	Patient satisfaction - Community Based Care	☐	☐	☑	☐	☐	☑	☐	☐	☐	☑	☑	☐	☐
PAS	Patient satisfaction - Health Care Services	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐
SCP	Physical activity - Stages of change	☑	☐	☑	☐	☑	☐	☐	☐	☐	☑	☐	☑	☐
PCU	Physical Check-up	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐
PSA	Prostate cancer screening	☐	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☑	☐
RAC	Restriction of Activities	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☑	☑
SWL	Satisfaction with life	☐	☐	☐	☐	☐	☐	☐	☐	☑	☑	☐	☐	☐
SFE	Self-esteem	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐
SXB	Sexual behaviours	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☑	☑
SLP	Sleep	☐	☐	☑	☐	☑	☐	☑	☐	☑	☐	☑	☐	☐
TAL	Smoking - Other tobacco products	☐	☐	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☐
SPC	Smoking - Physician Counselling	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑

Description	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt
SCH Smoking - Stages of Change	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑	☑	☐	☐
YSM Smoking - Youth smoking	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐
SCA Smoking cessation methods	☐	☐	☑	☐	☐	☐	☐	☐	☐	☑	☐	☐	☑
SPS Social Provisions	☐	☐	☐	☐	☑	☐	☐	☐	☐	☑	☑	☑	☑
SSA Social Support - Availability	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐
SPR Spiritual Values	☐	☐	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☐
SPI Spirometry	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
CST Stress - Child and adult stressors	☐	☐	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐	☐
RLE Stress - Recent life events	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☑	☐	☐
STS Stress - Sources	☐	☐	☐	☐	☑	☐	☐	☐	☐	☐	☑	☐	☐
SUI Suicidal thoughts and attempts	☑	☐	☐	☐	☐	☐	☐	☑	☑	☐	☐	☐	☑
SSB Sun safety behaviours	☐	☐	☐	☑	☐	☐	☐	☐	☐	☐	☐	☐	☐
UCN Unmet health care needs	☐	☐	☑	☐	☐	☑	☐	☐	☐	☑	☐	☑	☑
UPE Use of Protective Equipment	☑	☑	☐	☐	☐	☑	☐	☐	☑	☐	☑	☐	☐
ORG Voluntary organizations - Participation	☐	☑	☐	☐	☑	☐	☐	☐	☐	☐	☐	☑	☐

Appendix C - Available geography in the master and share files and their corresponding codes: Canada, provinces/territories, health regions and peer groups

Appendix C - Available geography in the master and share files and their corresponding codes: Canada, provinces/territories, health regions and peer groups

0	Canada
10	Newfoundland and Labrador
1011-C	Eastern Regional Integrated Health Authority
1012-I	Central Regional Integrated Health Authority
1013-I	Western Regional Integrated Health Authority
1014-H	Labrador-Grenfell Regional Integrated Health Authority
11	Prince Edward Island
1100-A	Prince Edward Island
12	Nova Scotia
1210-C	South Shore -- South West Nova
1223-C	Annapolis Valley District Health Authority
1230-C	Colchester East Hants -- Cumberland
1240-C	Pictou County -- Guysborough Antigonish Strait
1258-I	Cape Breton District Health Authority
1269-A	Capital District Health Authority
13	New Brunswick
1301-C	Zone 1 (Moncton area)
1302-C	Zone 2 (Saint John area)
1303-C	Zone 3 (Fredericton area)
1304-C	Zone 4 (Edmunston area)
1305-I	Zone 5 (Campbellton area)
1306-I	Zone 6 (Bathurst area)
1307-I	Zone 7 (Miramichi area)
24	Quebec
2401-C	Région du Bas-Saint-Laurent
2402-C	Région du Saguenay - Lac-Saint-Jean
2403-A	Région de la Capitale-Nationale
2404-C	Région de la Mauricie et du Centre-du-Québec
2405-A	Région de l'Estrie
2406-G	Région de Montréal
2407-A	Région de l'Outaouais
2408-C	Région de l'Abitibi-Témiscamingue
2409-H	Région de la Côte-Nord
2410-H	Région du Nord-du-Québec
2411-I	Région de la Gaspésie - Îles-de-la-Madeleine
2412-D	Région de la Chaudière-Appalaches
2413-A	Région de Laval
2414-A	Région de Lanaudière
2415-A	Région des Laurentides
2416-A	Région de la Montérégie
35	Ontario by Local Health Integration Network
3501	Erie St. Clair Health Integration Network
3502	South West Health Integration Network
3503	Waterloo Wellington Health Integration Network
3504	Hamilton Niagara Haldimand Brant Health Integration Network
3505	Central West Health Integration Network

3506	Mississauga Halton Health Integration Network
3507	Toronto Central Health Integration Network
3508	Central Health Integration Network
3509	Central East Health Integration Network
3510	South East Health Integration Network
3511	Champlain Health Integration Network
3512	North Simcoe Muskoka Health Integration Network
3513	North East Health Integration Network
3514	North West Health Integration Network
35	Ontario by Health Unit
3526-C	District of Algoma Health Unit
3527-A	Brant County Health Unit
3530-B	Durham Regional Health Unit
3531-A	Elgin-St. Thomas Health Unit
3533-D	Grey Bruce Health Unit
3534-A	Haldimand-Norfolk Health Unit
3535-A	Haliburton, Kawartha, Pine Ridge District Health Unit
3536-B	Halton Regional Health Unit
3537-A	City of Hamilton Health Unit
3538-A	Hastings and Prince Edward Counties Health Unit
3539-D	Huron County Health Unit
3540-C	Chatham-Kent Health Unit
3541-A	Kingston, Frontenac and Lennox and Addington Health Unit
3542-A	Lambton Health Unit
3543-A	Leeds, Grenville and Lanark District Health Unit
3544-A	Middlesex-London Health Unit
3546-A	Niagara Regional Area Health Unit
3547-C	North Bay Parry Sound District Health Unit
3549-H	Northwestern Health Unit
3551-B	City of Ottawa Health Unit
3552-A	Oxford County Health Unit
3553-J	Peel Regional Health Unit
3554-D	Perth District Health Unit
3555-A	Peterborough County-City Health Unit
3556-C	Porcupine Health Unit
3557-D	Renfrew County and District Health Unit
3558-A	Eastern Ontario Health Unit
3560-B	Simcoe Muskoka District Health Unit
3561-C	Sudbury and District Health Unit
3562-C	Thunder Bay District Health Unit
3563-C	Timiskaming Health Unit
3565-B	Waterloo Health Unit
3566-B	Wellington-Dufferin-Guelph Health Unit
3568-A	Windsor-Essex County Health Unit
3570-J	York Regional Health Unit
3595-G	City of Toronto Health Unit

46	Manitoba	
4601-A		Winnipeg Regional Health Authority
4602-D		Prairie Mountain Health
4603-E		Interlake-Eastern Regional Health Authority
4604-F		Northern Regional Health Authority
4605-D		Southern Health
47	Saskatchewan	
4701-D		Sun Country Regional Health Authority
4702-D		Five Hills Regional Health Authority
4703-D		Cypress Regional Health Authority
4704-A		Regina Qu'Appelle Regional Health Authority
4705-D		Sunrise Regional Health Authority
4706-A		Saskatoon Regional Health Authority
4707-D		Heartland Regional Health Authority
4708-D		Kelsey Trail Regional Health Authority
4709-H		Prince Albert Parkland Regional Health Authority
4710-H		Prairie North Regional Health Authority
4714-F		Mamawetan/Keewatin/Athabasca
48	Alberta	
4831-A		South Zone
4832-B		Calgary Zone
4833-E		Central Zone
4834-B		Edmonton Zone
4835-E		North Zone
59	British Columbia	
5911-D		East Kootenay Health Service Delivery Area
5912-D		Kootenay-Boundary Health Service Delivery Area
5913-A		Okanagan Health Service Delivery Area
5914-A		Thompson/Cariboo Health Service Delivery Area
5921-A		Fraser East Health Service Delivery Area
5922-J		Fraser North Health Service Delivery Area
5923-J		Fraser South Health Service Delivery Area
5931-J		Richmond Health Service Delivery Area
5932-G		Vancouver Health Service Delivery Area
5933-J		North Shore/Coast Garibaldi Health Service Delivery Area
5941-A		South Vancouver Island Health Service Delivery Area
5942-A		Central Vancouver Island Health Service Delivery Area
5943-A		North Vancouver Island Health Service Delivery Area
5951-H		Northwest Health Service Delivery Area
5952-H		Northern Interior Health Service Delivery Area
5953-E		Northeast Health Service Delivery Area
60	Yukon	
6001-E		Yukon
61	Northwest Territories	
6101-E		Northwest Territories
62	Nunavut – 10 largest communities	
6201-F		Nunavut – 10 largest communities
A	Peer group A	

B	Peer group B
C	Peer group C
D	Peer group D
E	Peer group E
F	Peer group F
G	Peer group G
H	Peer group H
I	Peer group I
J	Peer group J

Appendix D (2012) Sample allocation by health region and frame, and sample allocation by Local Health Integrated Network (LHIN) and frame in Ontario

Appendix D (2012) Sample allocation by health region and frame[16]

Geography		Area Frame		Phone frames		Combined	
Prov./Terr.	Health Region	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
Canada	**Total**	27,394	43,421	38,354	71,226	65,748	114,647
N.L.	**Total**	822	1,239	1,183	2,019	2,005	3,258
	1011	332	484	478	800	810	1,284
	1012	193	314	277	482	470	796
	1013	174	258	251	395	425	653
	1014	123	183	177	342	300	525
P.E.I.	**Total**	411	709	590	1,201	1,001	1,910
	1101	73	134	105	273	178	407
	1102	189	354	271	512	460	866
	1103	149	221	214	416	363	637
N.S.	**Total**	1,034	1,692	1,487	2,481	2,521	4,173
	1210	162	248	234	398	396	646
	1223	131	196	189	303	320	499
	1230	148	254	212	334	360	588
	1240	144	258	206	358	350	616
	1258	172	266	248	418	420	684
	1269	277	470	398	670	675	1,140
N.B.	**Total**	1,057	1,700	1,518	2,568	2,575	4,268
	1301	205	356	295	512	500	868
	1302	199	344	286	458	485	802
	1303	193	302	277	446	470	748
	1304	111	183	159	276	270	459
	1305	103	140	147	252	250	392
	1306	141	195	204	348	345	543
	1307	105	180	150	276	255	456
Que.	**Total**	4,813	7,334	7,333	14,202	12,146	21,536
	2401	246	366	354	630	600	996
	2402	257	365	371	658	628	1,023
	2403	378	561	548	931	926	1,492
	2404	329	479	474	786	803	1,265
	2405	253	440	365	604	618	1,044
	2406	637	1,064	916	1,724	1,553	2,788
	2407	264	408	381	651	645	1,059
	2408	246	366	354	582	600	948
	2409	246	354	354	792	600	1,146
	2410	0	0	400	1,838	400	1,838
	2411	246	348	354	714	600	1,062
	2412	296	424	427	758	723	1,182
	2413	275	419	395	730	670	1,149
	2414	294	462	424	734	718	1,196

[16] As mentioned in section 5.2, the figures for Manitoba and Prince Edward Island are based on the definitions of health regions that were used at the time of sampling.

Geography		Area Frame		Phone frames		Combined	
Prov./Terr.	Health Region	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
	2415	313	512	449	811	762	1,323
	2416	533	766	767	1,259	1,300	2,025
Ont.	**Total**	9,103	14,630	13,095	24,236	22,198	38,866
	3526	174	246	251	454	425	700
	3527	166	260	239	412	405	672
	3530	334	512	481	854	815	1,366
	3531	139	182	201	360	340	542
	3533	205	335	296	585	501	920
	3534	159	257	227	406	386	663
	3535	195	384	280	560	475	944
	3536	289	457	416	734	705	1,191
	3537	338	580	487	932	825	1,512
	3538	193	307	277	500	470	807
	3539	121	176	174	360	295	536
	3540	164	218	236	406	400	624
	3541	207	387	298	524	505	911
	3542	178	280	257	496	435	776
	3543	195	293	280	478	475	771
	3544	308	482	442	788	750	1,270
	3546	314	458	451	787	765	1,245
	3547	164	292	236	502	400	794
	3549	161	311	233	508	394	819
	3551	420	714	605	1,042	1,025	1,756
	3552	154	224	221	353	375	577
	3553	544	856	781	1,520	1,325	2,376
	3554	133	187	192	318	325	505
	3555	174	270	251	508	425	778
	3556	154	236	221	388	375	624
	3557	154	248	221	376	375	624
	3558	213	318	307	511	520	829
	3560	452	805	651	1,278	1,103	2,083
	3561	221	358	319	577	540	935
	3562	265	488	381	840	646	1,328
	3563	103	182	147	258	250	440
	3565	314	464	451	746	765	1,210
	3566	239	369	344	548	583	917
	3568	293	478	422	802	715	1,280
	3570	387	611	556	1,063	943	1,674
	3595	879	1,405	1,263	2,462	2,142	3,867
Man.	**Total**	1,540	2,396	2,210	3,985	3,750	6,381
	4610	433	715	622	970	1,055	1,685
	4615	115	163	165	264	280	427
	4620	103	163	147	294	250	457
	4625	123	168	177	291	300	459
	4630	141	258	204	396	345	654
	4640	164	230	236	404	400	634

Geography		Area Frame		Phone frames		Combined	
Prov./Terr.	Health Region	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
	4645	146	214	209	370	355	584
	4660	109	175	156	270	265	445
	4670	103	164	147	360	250	524
	4685	103	146	147	366	250	512
Sask.	**Total**	1,481	2,244	2,379	4,863	3,860	7,107
	4701	123	165	177	282	300	447
	4702	123	180	177	330	300	510
	4703	109	163	156	282	265	445
	4704	254	358	366	630	620	988
	4705	127	187	183	294	310	481
	4706	271	452	389	664	660	1,116
	4707	111	180	159	258	270	438
	4708	107	161	153	255	260	416
	4709	133	212	192	372	325	584
	4710	123	186	177	348	300	534
	4714	0	0	250	1,148	250	1,148
Alta.	**Total**	2,500	3,936	3,597	6,408	6,097	10,344
	4831	379	562	546	907	925	1,469
	4832	571	831	823	1,496	1,394	2,327
	4833	469	764	675	1,224	1,144	1,988
	4834	537	877	772	1,303	1,309	2,180
	4835	544	902	781	1,478	1,325	2,380
B.C.	**Total**	3,298	5,320	4,747	8,761	8,045	14,081
	5911	125	190	180	330	305	520
	5912	127	188	183	327	310	515
	5913	242	326	348	606	590	932
	5914	205	338	295	524	500	862
	5921	213	345	307	572	520	917
	5922	312	450	448	824	760	1,274
	5923	328	496	472	898	800	1,394
	5931	174	258	251	538	425	796
	5932	328	566	472	974	800	1,540
	5933	223	500	322	626	545	1,126
	5941	277	422	398	668	675	1,090
	5942	215	322	310	502	525	824
	5943	109	175	156	270	265	445
	5951	133	260	192	354	325	614
	5952	174	306	251	454	425	760
	5953	113	178	162	294	275	472
Y.T.	**6001**	475	744	125	256	600	1,000
N.W.T.	**6101**	510	855	90	246	600	1,101
Nvt.	**6201**	350	622	0	0	350	622

Appendix D (2012) - Sample allocation by Local Health Integrated Network (LHIN) and frame in Ontario

Geography		Area Frame		Phone frames		Combined	
Province	LHIN	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
Ont.	Total	9,103	14,630	13,095	24,236	22,198	38,866
	3501	635	976	915	1,704	1,550	2,680
	3502	1,050	1,545	1,511	2,723	2,561	4,268
	3503	510	777	733	1,202	1,243	1,979
	3504	1,065	1,689	1,532	2,745	2,597	4,434
	3505	431	716	618	1,187	1,049	1,903
	3506	458	699	657	1,246	1,115	1,945
	3507	461	725	662	1,306	1,123	2,031
	3508	578	908	833	1,576	1,411	2,484
	3509	864	1,429	1,244	2,364	2,108	3,793
	3510	539	899	774	1,373	1,313	2,272
	3511	843	1,368	1,214	2,058	2,057	3,426
	3512	427	786	614	1,225	1,041	2,011
	3513	816	1,314	1,174	2,179	1,990	3,493
	3514	426	799	614	1,348	1,040	2,147

Appendix E (2012) - Response rates by health region and frame, and response rates by Local Health Integrated Network (LHIN) and frame in Ontario

Appendix E (2012) - Table 9.1 response rates by health region and frame

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
Canada	**Total**	36,175	29,396	81.3	29,396	26,154	89.0	72.3	56,507	42,218	74.7	42,218	35,949	85.2	63.6	67.0
N.L.	**Total**	970	835	86.1	835	739	88.5	76.2	1,652	1,245	75.4	1,245	1,026	82.4	62.1	67.3
	1011	376	298	79.3	298	264	88.6	70.2	679	505	74.4	505	418	82.8	61.6	64.6
	1012	249	227	91.2	227	199	87.7	79.9	381	290	76.1	290	241	83.1	63.3	69.8
	1013	197	176	89.3	176	158	89.8	80.2	313	250	79.9	250	202	80.8	64.5	70.6
	1014	148	134	90.5	134	118	88.1	79.7	279	200	71.7	200	165	82.5	59.1	66.3
P.E.I.	**Total**	538	441	82.0	441	386	87.5	71.7	797	585	73.4	585	497	85.0	62.4	66.1
	1100	538	441	82.0	441	386	87.5	71.7	797	585	73.4	585	497	85.0	62.4	66.1
N.S.	**Total**	1,323	1,089	82.3	1,089	940	86.3	71.1	1,867	1,486	79.6	1,486	1,294	87.1	69.3	70.0
	1210	175	156	89.1	156	141	90.4	80.6	302	236	78.1	236	210	89.0	69.5	73.6
	1223	152	118	77.6	118	95	80.5	62.5	227	185	81.5	185	158	85.4	69.6	66.8
	1230	186	163	87.6	163	151	92.6	81.2	241	193	80.1	193	165	85.5	68.5	74.0
	1240	198	159	80.3	159	141	88.7	71.2	281	224	79.7	224	194	86.6	69.0	69.9
	1258	207	161	77.8	161	140	87.0	67.6	313	240	76.7	240	200	83.3	63.9	65.4
	1269	405	332	82.0	332	272	81.9	67.2	503	408	81.1	408	367	90.0	73.0	70.4
N.B.	**Total**	1,368	1,130	82.6	1,130	1,007	89.1	73.6	2,081	1,666	80.1	1,666	1,450	87.0	69.7	71.2
	1301	292	224	76.7	224	205	91.5	70.2	429	343	80.0	343	302	88.0	70.4	70.3
	1302	261	202	77.4	202	180	89.1	69.0	382	314	82.2	314	274	87.3	71.7	70.6
	1303	250	205	82.0	205	180	87.8	72.0	352	285	81.0	285	255	89.5	72.4	72.3
	1304	151	139	92.1	139	131	94.2	86.8	229	172	75.1	172	148	86.0	64.6	73.4
	1305	113	97	85.8	97	84	86.6	74.3	195	163	83.6	163	139	85.3	71.3	72.4
	1306	156	143	91.7	143	130	90.9	83.3	292	224	76.7	224	191	85.3	65.4	71.7
	1307	145	120	82.8	120	97	80.8	66.9	202	165	81.7	165	141	85.5	69.8	68.6
Que.	**Total**	6,277	5,198	82.8	5,198	4,785	92.1	76.2	11,177	8,394	75.1	8,394	7,184	85.6	64.3	68.6
	2401	279	250	89.6	250	233	93.2	83.5	550	430	78.2	430	380	88.4	69.1	73.9
	2402	326	294	90.2	294	276	93.9	84.7	565	431	76.3	431	376	87.2	66.5	73.2
	2403	515	419	81.4	419	386	92.1	75.0	813	615	75.6	615	524	85.2	64.5	68.5

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	2404	407	350	86.0	350	324	92.6	79.6	691	546	79.0	546	486	89.0	70.3	73.8
	2405	348	281	80.7	281	267	95.0	76.7	496	396	79.8	396	353	89.1	71.2	73.5
	2406	933	708	75.9	708	655	92.5	70.2	1,498	1,037	69.2	1,037	864	83.3	57.7	62.5
	2407	319	266	83.4	266	245	92.1	76.8	527	410	77.8	410	353	86.1	67.0	70.7
	2408	296	240	81.1	240	221	92.1	74.7	480	385	80.2	385	334	86.8	69.6	71.5
	2409	298	254	85.2	254	236	92.9	79.2	595	417	70.1	417	357	85.6	60.0	66.4
	2410	.	.	.	.	.	.	.	609	483	79.3	483	396	82.0	65.0	65.0
	2411	255	212	83.1	212	195	92.0	76.5	581	417	71.8	417	348	83.5	59.9	65.0
	2412	354	302	85.3	302	279	92.4	78.8	689	518	75.2	518	446	86.1	64.7	69.5
	2413	394	312	79.2	312	273	87.5	69.3	649	466	71.8	466	388	83.3	59.8	63.4
	2414	377	326	86.5	326	298	91.4	79.0	626	468	74.8	468	391	83.5	62.5	68.7
	2415	480	407	84.8	407	374	91.9	77.9	674	496	73.6	496	428	86.3	63.5	69.5
	2416	696	577	82.9	577	523	90.6	75.1	1,134	879	77.5	879	760	86.5	67.0	70.1
Ont.	**Total**	12,303	9,778	79.5	9,778	8,646	88.4	70.3	19,895	14,968	75.2	14,968	12,611	84.3	63.4	66.0
	3526	208	158	76.0	158	148	93.7	71.2	360	273	75.8	273	243	89.0	67.5	68.8
	3527	240	197	82.1	197	175	88.8	72.9	352	264	75.0	264	232	87.9	65.9	68.8
	3530	447	353	79.0	353	303	85.8	67.8	719	559	77.7	559	453	81.0	63.0	64.8
	3531	155	132	85.2	132	124	93.9	80.0	286	228	79.7	228	190	83.3	66.4	71.2
	3533	219	192	87.7	192	176	91.7	80.4	442	330	74.7	330	288	87.3	65.2	70.2
	3534	204	150	73.5	150	125	83.3	61.3	331	254	76.7	254	215	84.6	65.0	63.6
	3535	275	227	82.5	227	201	88.5	73.1	430	334	77.7	334	296	88.6	68.8	70.5
	3536	410	348	84.9	348	311	89.4	75.9	656	496	75.6	496	406	81.9	61.9	67.3
	3537	486	359	73.9	359	307	85.5	63.2	785	580	73.9	580	477	82.2	60.8	61.7
	3538	240	210	87.5	210	179	85.2	74.6	380	299	78.7	299	264	88.3	69.5	71.5
	3539	146	129	88.4	129	120	93.0	82.2	271	209	77.1	209	183	87.6	67.5	72.7
	3540	180	163	90.6	163	156	95.7	86.7	337	258	76.6	258	220	85.3	65.3	72.7
	3541	305	220	72.1	220	180	81.8	59.0	423	323	76.4	323	288	89.2	68.1	64.3
	3542	237	202	85.2	202	190	94.1	80.2	411	317	77.1	317	270	85.2	65.7	71.0
	3543	258	202	78.3	202	171	84.7	66.3	398	318	79.9	318	290	91.2	72.9	70.3
	3544	393	327	83.2	327	305	93.3	77.6	709	525	74.0	525	438	83.4	61.8	67.4

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	3546	422	341	80.8	341	300	88.0	71.1	659	499	75.7	499	424	85.0	64.3	67.0
	3547	185	155	83.8	155	139	89.7	75.1	349	272	77.9	272	238	87.5	68.2	70.6
	3549	244	196	80.3	196	176	89.8	72.1	319	244	76.5	244	214	87.7	67.1	69.3
	3551	656	500	76.2	500	438	87.6	66.8	909	728	80.1	728	614	84.3	67.5	67.2
	3552	186	171	91.9	171	163	95.3	87.6	309	251	81.2	251	216	86.1	69.9	76.6
	3553	809	662	81.8	662	599	90.5	74.0	1,306	943	72.2	943	740	78.5	56.7	63.3
	3554	163	143	87.7	143	129	90.2	79.1	279	230	82.4	230	201	87.4	72.0	74.7
	3555	204	166	81.4	166	145	87.3	71.1	362	280	77.3	280	250	89.3	69.1	69.8
	3556	185	141	76.2	141	129	91.5	69.7	313	241	77.0	241	208	86.3	66.5	67.7
	3557	167	136	81.4	136	123	90.4	73.7	294	241	82.0	241	212	88.0	72.1	72.7
	3558	283	236	83.4	236	211	89.4	74.6	428	359	83.9	359	317	88.3	74.1	74.3
	3560	640	478	74.7	478	418	87.4	65.3	960	740	77.1	740	648	87.6	67.5	66.6
	3561	288	230	79.9	230	201	87.4	69.8	439	353	80.4	353	308	87.3	70.2	70.0
	3562	389	283	72.8	283	233	82.3	59.9	652	508	77.9	508	443	87.2	67.9	64.9
	3563	121	86	71.1	86	80	93.0	66.1	210	166	79.0	166	135	81.3	64.3	65.0
	3565	423	341	80.6	341	301	88.3	71.2	666	501	75.2	501	414	82.6	62.2	65.7
	3566	313	272	86.9	272	247	90.8	78.9	476	401	84.2	401	337	84.0	70.8	74.0
	3568	399	301	75.4	301	264	87.7	66.2	682	513	75.2	513	424	82.7	62.2	63.6
	3570	572	459	80.2	459	377	82.1	65.9	934	608	65.1	608	464	76.3	49.7	55.8
	3595	1,251	912	72.9	912	802	87.9	64.1	2,059	1,323	64.3	1,323	1,051	79.4	51.0	56.0
Man.	**Total**	1,971	1,646	83.5	1,646	1,472	89.4	74.7	2,928	2,297	78.4	2,297	2,014	87.7	68.8	71.2
	4601	651	511	78.5	511	457	89.4	70.2	832	646	77.6	646	566	87.6	68.0	69.0
	4602	437	365	83.5	365	332	91.0	76.0	707	542	76.7	542	480	88.6	67.9	71.0
	4603	313	284	90.7	284	253	89.1	80.8	448	351	78.3	351	309	88.0	69.0	73.9
	4604	234	194	82.9	194	171	88.1	73.1	393	310	78.9	310	260	83.9	66.2	68.7
	4605	336	292	86.9	292	259	88.7	77.1	548	448	81.8	448	399	89.1	72.8	74.4
Sask.	**Total**	1,814	1,540	84.9	1,540	1,386	90.0	76.4	3,342	2,484	74.3	2,484	2,206	88.8	66.0	69.7
	4701	132	117	88.6	117	114	97.4	86.4	232	179	77.2	179	161	89.9	69.4	75.5
	4702	147	127	86.4	127	116	91.3	78.9	266	198	74.4	198	176	88.9	66.2	70.7

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	4703	123	111	90.2	111	100	90.1	81.3	231	170	73.6	170	146	85.9	63.2	69.5
	4704	298	255	85.6	255	234	91.8	78.5	520	382	73.5	382	333	87.2	64.0	69.3
	4705	151	144	95.4	144	135	93.8	89.4	238	178	74.8	178	164	92.1	68.9	76.9
	4706	374	284	75.9	284	244	85.9	65.2	581	447	76.9	447	390	87.2	67.1	66.4
	4707	144	120	83.3	120	107	89.2	74.3	207	157	75.8	157	137	87.3	66.2	69.5
	4708	119	105	88.2	105	96	91.4	80.7	198	157	79.3	157	143	91.1	72.2	75.4
	4709	169	135	79.9	135	110	81.5	65.1	259	204	78.8	204	191	93.6	73.7	70.3
	4710	157	142	90.4	142	130	91.5	82.8	262	185	70.6	185	172	93.0	65.6	72.1
	4714	.	.	.	.	.	.	.	348	227	65.2	227	193	85.0	55.5	55.5
Alta.	**Total**	3,406	2,790	81.9	2,790	2,438	87.4	71.6	5,329	3,727	69.9	3,727	3,082	82.7	57.8	63.2
	4831	454	398	87.7	398	352	88.4	77.5	747	512	68.5	512	439	85.7	58.8	65.9
	4832	797	623	78.2	623	535	85.9	67.1	1,304	929	71.2	929	754	81.2	57.8	61.4
	4833	657	534	81.3	534	471	88.2	71.7	987	708	71.7	708	599	84.6	60.7	65.1
	4834	769	621	80.8	621	547	88.1	71.1	1,121	778	69.4	778	635	81.6	56.6	62.5
	4835	729	614	84.2	614	533	86.8	73.1	1,170	800	68.4	800	655	81.9	56.0	62.6
B.C.	**Total**	4,447	3,478	78.2	3,478	3,088	88.8	69.4	7,215	5,189	71.9	5,189	4,429	85.4	61.4	64.5
	5911	157	131	83.4	131	112	85.5	71.3	271	198	73.1	198	173	87.4	63.8	66.6
	5912	156	129	82.7	129	114	88.4	73.1	280	214	76.4	214	187	87.4	66.8	69.0
	5913	271	238	87.8	238	225	94.5	83.0	507	364	71.8	364	313	86.0	61.7	69.2
	5914	278	214	77.0	214	200	93.5	71.9	418	305	73.0	305	257	84.3	61.5	65.7
	5921	305	244	80.0	244	216	88.5	70.8	456	341	74.8	341	291	85.3	63.8	66.6
	5922	407	302	74.2	302	266	88.1	65.4	707	509	72.0	509	415	81.5	58.7	61.1
	5923	442	369	83.5	369	325	88.1	73.5	755	526	69.7	526	429	81.6	56.8	63.0
	5931	222	194	87.4	194	178	91.8	80.2	455	311	68.4	311	258	83.0	56.7	64.4
	5932	476	347	72.9	347	321	92.5	67.4	773	500	64.7	500	418	83.6	54.1	59.2
	5933	364	307	84.3	307	259	84.4	71.2	514	344	66.9	344	300	87.2	58.4	63.7
	5941	369	289	78.3	289	265	91.7	71.8	558	426	76.3	426	378	88.7	67.7	69.4
	5942	262	216	82.4	216	192	88.9	73.3	413	322	78.0	322	284	88.2	68.8	70.5
	5943	133	85	63.9	85	83	97.6	62.4	232	175	75.4	175	148	84.6	63.8	63.3

Geography		Area frame							Phone frames							Combined
Prov. Terr.	**Health Region**	**# in scope HH**	**# resp. HH**	**HH resp. rates**	**# pers. select.**	**# resp.**	**Pers. resp. rates**	**Resp. rates**	**# in scope HH**	**# resp. HH**	**HH resp. rates**	**# pers. select.**	**# resp.**	**Pers. resp. rates**	**Resp. rates**	**Combined resp. rates**
	5951	216	168	77.8	168	146	86.9	67.6	254	177	69.7	177	153	86.4	60.2	63.6
	5952	251	155	61.8	155	118	76.1	47.0	382	293	76.7	293	264	90.1	69.1	60.3
	5953	138	90	65.2	90	68	75.6	49.3	240	184	76.7	184	161	87.5	67.1	60.6
Y.T.	**6001**	595	474	79.7	474	418	88.2	70.3	131	103	78.6	103	94	91.3	71.8	70.5
N.W.T.	**6101**	688	587	85.3	587	486	82.8	70.6	93	74	79.6	74	62	83.8	66.7	70.2
Nvt.	**6201**	475	410	86.3	410	363	88.5	76.4	.	.	.	.	.	.	.	76.4

Appendix E (2012) - Table 9.2 Response rate by Local Health Integrated Network (LHIN) and frame in Ontario

Geography		Area frame							Phone frames							Combined
Province	LHIN	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
Ont.	**Total**	12,303	9,778	79.5	9,778	8,646	88.4	70.3	19,895	14,968	75.2	14,968	12,611	84.3	63.4	66.0
	3501	816	666	81.6	666	610	91.6	74.8	1,430	1,088	76.1	1,088	914	84.0	63.9	67.9
	3502	1,246	1,079	86.6	1,079	1,003	93.0	80.5	2,265	1,758	77.6	1,758	1,497	85.2	66.1	71.2
	3503	704	594	84.4	594	530	89.2	75.3	1,067	831	77.9	831	694	83.5	65.0	69.1
	3504	1,495	1,158	77.5	1,158	1,001	86.4	67.0	2,327	1,749	75.2	1,749	1,468	83.9	63.1	64.6
	3505	629	502	79.8	502	455	90.6	72.3	998	707	70.8	707	566	80.1	56.7	62.8
	3506	662	560	84.6	560	509	90.9	76.9	1,097	809	73.7	809	641	79.2	58.4	65.4
	3507	625	451	72.2	451	398	88.2	63.7	1,069	695	65.0	695	555	79.9	51.9	56.3
	3508	870	689	79.2	689	572	83.0	65.7	1,368	897	65.6	897	696	77.6	50.9	56.7
	3509	1,141	888	77.8	888	773	87.0	67.7	1,879	1,401	74.6	1,401	1,182	84.4	62.9	64.7
	3510	743	594	79.9	594	498	83.8	67.0	1,096	862	78.6	862	773	89.7	70.5	69.1
	3511	1,188	932	78.5	932	825	88.5	69.4	1,746	1,416	81.1	1,416	1,222	86.3	70.0	69.8
	3512	564	416	73.8	416	366	88.0	64.9	910	697	76.6	697	613	87.9	67.4	66.4
	3513	987	770	78.0	770	697	90.5	70.6	1,672	1,306	78.1	1,306	1,133	86.8	67.8	68.8
	3514	633	479	75.7	479	409	85.4	64.6	971	752	77.4	752	657	87.4	67.7	66.5

Appendix F (2011-2012) - Sample allocation by health region and frame and sample allocation by Local Health Integrated Network (LHIN) and frame in Ontario

Appendix F (2011-2012) Sample allocation by health region and frame[17]

Geography		Area Frame		Phone frames		Combined	
Prov./Terr.	Health Region	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
Canada	**Total**	54,860	87,320	76,638	139,681	131,498	227,001
N.L.	**Total**	1,646	2,539	2,364	3,947	4,010	6,486
	1011	665	1,043	955	1,593	1,620	2,636
	1012	386	625	554	924	940	1,549
	1013	349	506	501	799	850	1,305
	1014	246	365	354	631	600	996
P.E.I.	**Total**	822	1,425	1,180	2,397	2,002	3,822
	1101	146	264	210	517	356	781
	1102	378	674	542	1,048	920	1,722
	1103	298	487	428	832	726	1,319
N.S.	**Total**	2,071	3,391	2,971	4,880	5,042	8,271
	1210	324	506	468	780	792	1,286
	1223	263	408	377	605	640	1,013
	1230	296	496	424	689	720	1,185
	1240	288	489	412	688	700	1,177
	1258	345	542	495	840	840	1,382
	1269	555	950	795	1,278	1,350	2,228
N.B.	**Total**	2,116	3,508	3,034	4,919	5,150	8,427
	1301	411	696	589	958	1,000	1,654
	1302	398	704	572	924	970	1,628
	1303	386	662	554	859	940	1,521
	1304	222	362	318	514	540	876
	1305	206	325	294	468	500	793
	1306	283	391	407	670	690	1,061
	1307	210	368	300	526	510	894
Que.	**Total**	9,644	14,713	14,647	27,612	24,291	42,325
	2401	493	711	707	1,224	1,200	1,935
	2402	515	747	741	1,218	1,256	1,965
	2403	760	1,159	1,092	1,814	1,852	2,973
	2404	659	951	947	1,529	1,606	2,480
	2405	507	894	729	1,186	1,236	2,080
	2406	1,276	2,110	1,830	3,456	3,106	5,566
	2407	529	839	761	1,276	1,290	2,115
	2408	493	726	707	1,131	1,200	1,857
	2409	493	702	707	1,548	1,200	2,250
	2410	0	0	800	3,466	800	3,466
	2411	493	697	707	1,398	1,200	2,095
	2412	593	820	853	1,486	1,446	2,306
	2413	550	856	790	1,396	1,340	2,252
	2414	589	897	847	1,427	1,436	2,324

[17] As mentioned in section 5.2, the figures for Manitoba and Prince Edward Island are based on the definitions of HRs that were used at the time of sampling.

Geography		Area Frame		Phone frames		Combined	
Prov./Terr.	Health Region	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
	2415	626	1,032	897	1,595	1,523	2,627
	2416	1,068	1,572	1,532	2,462	2,600	4,034
Ont.	**Total**	18,228	29,090	26,168	47,515	44,396	76,605
	3526	349	492	501	892	850	1,384
	3527	333	530	477	810	810	1,340
	3530	669	1,019	961	1,685	1,630	2,704
	3531	279	408	401	684	680	1,092
	3533	411	635	591	1,167	1,002	1,802
	3534	318	518	454	830	772	1,348
	3535	390	780	560	1,102	950	1,882
	3536	579	884	831	1,434	1,410	2,318
	3537	677	1,150	973	1,752	1,650	2,902
	3538	386	598	554	997	940	1,595
	3539	242	345	348	664	590	1,009
	3540	328	437	472	816	800	1,253
	3541	415	722	595	1,057	1,010	1,779
	3542	357	551	513	904	870	1,455
	3543	390	635	560	975	950	1,610
	3544	616	974	884	1,565	1,500	2,539
	3546	629	935	901	1,591	1,530	2,526
	3547	328	566	472	1,008	800	1,574
	3549	323	640	465	976	788	1,616
	3551	841	1,438	1,209	2,031	2,050	3,469
	3552	308	427	442	711	750	1,138
	3553	1,088	1,664	1,562	2,860	2,650	4,524
	3554	267	363	383	632	650	995
	3555	349	572	501	969	850	1,541
	3556	308	482	442	796	750	1,278
	3557	308	486	442	778	750	1,264
	3558	427	648	613	1,039	1,040	1,687
	3560	905	1,603	1,301	2,532	2,206	4,135
	3561	443	738	637	1,112	1,080	1,850
	3562	530	932	762	1,556	1,292	2,488
	3563	206	372	294	524	500	896
	3565	629	941	901	1,488	1,530	2,429
	3566	478	688	688	1,120	1,166	1,808
	3568	587	917	843	1,554	1,430	2,471
	3570	775	1,187	1,111	2,027	1,886	3,214
	3595	1,760	2,813	2,524	4,877	4,284	7,690
Man.	**Total**	3,082	4,847	4,418	7,768	7,500	12,615
	4610	867	1,411	1,243	1,927	2,110	3,338
	4615	230	354	330	537	560	891
	4620	206	320	294	572	500	892
	4625	246	366	354	589	600	955
	4630	283	514	407	764	690	1,278

Geography		Area Frame		Phone frames		Combined	
Prov./Terr.	Health Region	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
	4640	328	454	472	826	800	1,280
	4645	292	420	418	719	710	1,139
	4660	218	366	312	532	530	898
	4670	206	322	294	642	500	964
	4685	206	320	294	660	500	980
Sask.	**Total**	2,964	4,440	4,756	9,420	7,720	13,860
	4701	246	332	354	569	600	901
	4702	246	357	354	616	600	973
	4703	218	333	312	532	530	865
	4704	509	709	731	1,226	1,240	1,935
	4705	254	384	366	617	620	1,001
	4706	542	870	778	1,296	1,320	2,166
	4707	222	364	318	514	540	878
	4708	214	318	306	480	520	798
	4709	267	400	383	714	650	1,114
	4710	246	373	354	674	600	1,047
	4714	0	0	500	2,182	500	2,182
Alta.	**Total**	5,010	8,126	7,187	12,691	12,197	20,817
	4831	760	1,143	1,090	1,833	1,850	2,976
	4832	1,145	1,748	1,644	2,930	2,789	4,678
	4833	940	1,545	1,349	2,400	2,289	3,945
	4834	1,076	1,863	1,543	2,642	2,619	4,505
	4835	1,089	1,827	1,561	2,886	2,650	4,713
B.C.	**Total**	6,607	10,773	9,483	17,414	16,090	28,187
	5911	250	385	360	634	610	1,019
	5912	254	404	366	643	620	1,047
	5913	485	666	695	1,199	1,180	1,865
	5914	411	646	589	1,056	1,000	1,702
	5921	427	672	613	1,138	1,040	1,810
	5922	624	922	896	1,642	1,520	2,564
	5923	657	1,009	943	1,740	1,600	2,749
	5931	349	519	501	1,034	850	1,553
	5932	657	1,144	943	1,960	1,600	3,104
	5933	447	996	643	1,246	1,090	2,242
	5941	555	876	795	1,394	1,350	2,270
	5942	431	652	619	1,013	1,050	1,665
	5943	218	360	312	510	530	870
	5951	267	484	383	714	650	1,198
	5952	349	630	501	895	850	1,525
	5903	226	408	324	596	550	1,004
Y.T.	**6001**	950	1,488	250	574	1,200	2,062
N.W.T.	**6101**	1,020	1,769	180	544	1,200	2,313
Nvt.	**6201**	700	1,211	0	0	700	1,211

Appendix F (2011-2012) - Sample allocation by Local Health Integrated Network (LHIN) and frame in Ontario

Geography		Area Frame		Phone frames		Combined	
Province	LHIN	expected # of respondents	raw sample size	expected # of respondents	raw sample size	expected # of respondents	raw sample size
Ont.	**Total**	18,228	29,090	26,168	47,515	44,396	76,605
	3501	1,272	1,905	1,828	3,274	3,100	5,179
	3502	2,103	3,087	3,019	5,348	5,122	8,435
	3503	1,021	1,521	1,465	2,394	2,486	3,915
	3504	2,134	3,395	3,060	5,412	5,194	8,807
	3505	862	1,379	1,236	2,301	2,098	3,680
	3506	916	1,357	1,314	2,360	2,230	3,717
	3507	922	1,469	1,324	2,550	2,246	4,019
	3508	1,158	1,790	1,664	3,058	2,822	4,848
	3509	1,731	2,882	2,485	4,638	4,216	7,520
	3510	1,079	1,781	1,547	2,773	2,626	4,554
	3511	1,688	2,746	2,426	4,104	4,114	6,850
	3512	855	1,556	1,227	2,439	2,082	3,995
	3513	1,634	2,650	2,346	4,332	3,980	6,982
	3514	853	1,572	1,227	2,532	2,080	4,104

Appendix G (2011-2012) - Response rates by health region and frame and response rates by Local Health Integrated Network (LHIN) and frame in Ontario

Appendix G (2011-2012) - Table 9.3 response rates by health region and frame

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
Canada	**Total**	72,808	59,361	81.5	59,361	53,165	89.6	73.0	110,913	84,639	76.3	84,639	72,480	85.6	65.3	68.4
N.L.	**Total**	2,011	1,738	86.4	1,738	1,561	89.8	77.6	3,230	2,470	76.5	2,470	2,064	83.6	63.9	69.2
	1011	843	687	81.5	687	611	88.9	72.5	1,329	1,006	75.7	1,006	844	83.9	63.5	67.0
	1012	483	438	90.7	438	386	88.1	79.9	737	572	77.6	572	478	83.6	64.9	70.8
	1013	391	347	88.7	347	319	91.9	81.6	649	509	78.4	509	423	83.1	65.2	71.3
	1014	294	266	90.5	266	245	92.1	83.3	515	383	74.4	383	319	83.3	61.9	69.7
P.E.I.	**Total**	1,053	879	83.5	879	767	87.3	72.8	1,583	1,197	75.6	1,197	1,007	84.1	63.6	67.3
	1101	1,053	879	83.5	879	767	87.3	72.8	1,583	1,197	75.6	1,197	1,007	84.1	63.6	67.3
N.S.	**Total**	2,705	2,241	82.8	2,241	1,959	87.4	72.4	3,745	3,054	81.5	3,054	2,670	87.4	71.3	71.8
	1210	372	338	90.9	338	306	90.5	82.3	600	478	79.7	478	424	88.7	70.7	75.1
	1223	326	271	83.1	271	233	86.0	71.5	457	383	83.8	383	333	86.9	72.9	72.3
	1230	380	323	85.0	323	295	91.3	77.6	499	410	82.2	410	364	88.8	72.9	75.0
	1240	374	305	81.6	305	267	87.5	71.4	542	440	81.2	440	382	86.8	70.5	70.9
	1258	439	360	82.0	360	314	87.2	71.5	631	501	79.4	501	426	85.0	67.5	69.2
	1269	814	644	79.1	644	544	84.5	66.8	1,016	842	82.9	842	741	88.0	72.9	70.2
N.B.	**Total**	2,810	2,310	82.2	2,310	2,050	88.7	73.0	3,950	3,166	80.2	3,166	2,736	86.4	69.3	70.8
	1301	576	434	75.3	434	388	89.4	67.4	783	618	78.9	618	542	87.7	69.2	68.4
	1302	547	421	77.0	421	364	86.5	66.5	764	619	81.0	619	548	88.5	71.7	69.6
	1303	513	434	84.6	434	389	89.6	75.8	694	569	82.0	569	503	88.4	72.5	73.9
	1304	302	269	89.1	269	254	94.4	84.1	412	320	77.7	320	272	85.0	66.0	73.7
	1305	256	213	83.2	213	187	87.8	73.0	353	296	83.9	296	252	85.1	71.4	72.1
	1306	321	295	91.9	295	270	91.5	84.1	543	413	76.1	413	335	81.1	61.7	70.0
	1307	295	244	82.7	244	198	81.1	67.1	401	331	82.5	331	284	85.8	70.8	69.3
Que.	**Total**	12,621	10,401	82.4	10,401	9,607	92.4	76.1	21,800	16,694	76.6	16,694	14,369	86.1	65.9	69.7
	2401	543	481	88.6	481	459	95.4	84.5	1,051	829	78.9	829	725	87.5	69.0	74.3

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	2402	667	596	89.4	596	563	94.5	84.4	1,041	822	79.0	822	728	88.6	69.9	75.6
	2403	1,055	842	79.8	842	778	92.4	73.7	1,598	1,196	74.8	1,196	1,031	86.2	64.5	68.2
	2404	796	677	85.1	677	625	92.3	78.5	1,342	1,089	81.1	1,089	965	88.6	71.9	74.4
	2405	730	585	80.1	585	557	95.2	76.3	973	792	81.4	792	705	89.0	72.5	74.1
	2406	1,864	1,400	75.1	1,400	1,289	92.1	69.2	3,019	2,169	71.8	2,169	1,817	83.8	60.2	63.6
	2407	665	560	84.2	560	512	91.4	77.0	1,028	805	78.3	805	705	87.6	68.6	71.9
	2408	585	493	84.3	493	452	91.7	77.3	957	788	82.3	788	691	87.7	72.2	74.1
	2409	571	491	86.0	491	460	93.7	80.6	1,191	856	71.9	856	724	84.6	60.8	67.2
	2410	.	.	.	.	.	.	.	1,072	855	79.8	855	716	83.7	66.8	66.8
	2411	548	483	88.1	483	448	92.8	81.8	1,152	841	73.0	841	703	83.6	61.0	67.7
	2412	684	574	83.9	574	533	92.9	77.9	1,362	1,062	78.0	1,062	911	85.8	66.9	70.6
	2413	811	637	78.5	637	560	87.9	69.1	1,234	916	74.2	916	770	84.1	62.4	65.0
	2414	755	642	85.0	642	591	92.1	78.3	1,213	930	76.7	930	791	85.1	65.2	70.2
	2415	929	768	82.7	768	704	91.7	75.8	1,322	986	74.6	986	843	85.5	63.8	68.7
	2416	1,418	1,172	82.7	1,172	1,076	91.8	75.9	2,245	1,758	78.3	1,758	1,544	87.8	68.8	71.5
Ont.	**Total**	24,559	19,559	79.6	19,559	17,367	88.8	70.7	38,865	30,035	77.3	30,035	25,548	85.1	65.7	67.7
	3526	417	335	80.3	335	316	94.3	75.8	688	552	80.2	552	496	89.9	72.1	73.5
	3527	476	376	79.0	376	324	86.2	68.1	659	509	77.2	509	442	86.8	67.1	67.5
	3530	899	724	80.5	724	635	87.7	70.6	1,449	1,150	79.4	1,150	948	82.4	65.4	67.4
	3531	348	287	82.5	287	267	93.0	76.7	553	440	79.6	440	379	86.1	68.5	71.7
	3533	460	392	85.2	392	358	91.3	77.8	878	688	78.4	688	608	88.4	69.2	72.2
	3534	418	312	74.6	312	263	84.3	62.9	657	516	78.5	516	437	84.7	66.5	65.1
	3535	530	413	77.9	413	363	87.9	68.5	809	635	78.5	635	559	88.0	69.1	68.9
	3536	804	665	82.7	665	598	89.9	74.4	1,272	998	78.5	998	844	84.6	66.4	69.5
	3537	973	727	74.7	727	626	86.1	64.3	1,504	1,137	75.6	1,137	954	83.9	63.4	63.8
	3538	481	414	86.1	414	350	84.5	72.8	766	613	80.0	613	544	88.7	71.0	71.7
	3539	281	251	89.3	251	234	93.2	83.3	500	396	79.2	396	346	87.4	69.2	74.3

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	3540	352	316	89.8	316	301	95.3	85.5	674	530	78.6	530	465	87.7	69.0	74.7
	3541	575	421	73.2	421	351	83.4	61.0	845	672	79.5	672	592	88.1	70.1	66.4
	3542	473	402	85.0	402	371	92.3	78.4	706	543	76.9	543	471	86.7	66.7	71.4
	3543	544	452	83.1	452	391	86.5	71.9	804	666	82.8	666	597	89.6	74.3	73.3
	3544	803	658	81.9	658	615	93.5	76.6	1,393	1,066	76.5	1,066	901	84.5	64.7	69.0
	3546	836	696	83.3	696	625	89.8	74.8	1,339	1,037	77.4	1,037	891	85.9	66.5	69.7
	3547	362	302	83.4	302	267	88.4	73.8	711	572	80.5	572	507	88.6	71.3	72.1
	3549	465	368	79.1	368	321	87.2	69.0	610	486	79.7	486	432	88.9	70.8	70.0
	3551	1,316	989	75.2	989	869	87.9	66.0	1,764	1,443	81.8	1,443	1,235	85.6	70.0	68.3
	3552	361	327	90.6	327	313	95.7	86.7	612	513	83.8	513	445	86.7	72.7	77.9
	3553	1,555	1,311	84.3	1,311	1,164	88.8	74.9	2,460	1,804	73.3	1,804	1,432	79.4	58.2	64.7
	3554	319	260	81.5	260	240	92.3	75.2	547	461	84.3	461	411	89.2	75.1	75.2
	3555	461	372	80.7	372	331	89.0	71.8	709	561	79.1	561	497	88.6	70.1	70.8
	3556	409	297	72.6	297	258	86.9	63.1	641	509	79.4	509	444	87.2	69.3	66.9
	3557	351	303	86.3	303	267	88.1	76.1	610	509	83.4	509	456	89.6	74.8	75.2
	3558	569	485	85.2	485	435	89.7	76.4	869	742	85.4	742	655	88.3	75.4	75.8
	3560	1,276	937	73.4	937	827	88.3	64.8	1,923	1,543	80.2	1,543	1,343	87.0	69.8	67.8
	3561	586	467	79.7	467	409	87.6	69.8	843	681	80.8	681	591	86.8	70.1	70.0
	3562	744	552	74.2	552	478	86.6	64.2	1,188	937	78.9	937	835	89.1	70.3	68.0
	3563	243	174	71.6	174	162	93.1	66.7	427	348	81.5	348	297	85.3	69.6	68.5
	3565	848	658	77.6	658	586	89.1	69.1	1,330	1,037	78.0	1,037	862	83.1	64.8	66.5
	3566	600	519	86.5	519	471	90.8	78.5	968	809	83.6	809	697	86.2	72.0	74.5
	3568	780	606	77.7	606	546	90.1	70.0	1,322	1,021	77.2	1,021	847	83.0	64.1	66.3
	3570	1,108	885	79.9	885	762	86.1	68.8	1,781	1,224	68.7	1,224	961	78.5	54.0	59.6
	3595	2,536	1,906	75.2	1,906	1,673	87.8	66.0	4,054	2,687	66.3	2,687	2,127	79.2	52.5	57.7
Man.	**Total**	4,035	3,354	83.1	3,354	3,003	89.5	74.4	5,750	4,494	78.2	4,494	3,959	88.1	68.9	71.1
	4601	1,295	1,011	78.1	1,011	895	88.5	69.1	1,668	1,311	78.6	1,311	1,144	87.3	68.6	68.8

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	4602	920	768	83.5	768	698	90.9	75.9	1,401	1,080	77.1	1,080	965	89.4	68.9	71.7
	4603	628	561	89.3	561	504	89.8	80.3	848	656	77.4	656	584	89.0	68.9	73.7
	4604	481	392	81.5	392	348	88.8	72.3	710	537	75.6	537	456	84.9	64.2	67.5
	4605	711	622	87.5	622	558	89.7	78.5	1,123	910	81.0	910	810	89.0	72.1	74.6
Sask.	**Total**	3,573	3,078	86.1	3,078	2,827	91.8	79.1	6,507	4,897	75.3	4,897	4,334	88.5	66.6	71.0
	4701	258	233	90.3	233	227	97.4	88.0	468	354	75.6	354	317	89.5	67.7	74.9
	4702	296	265	89.5	265	251	94.7	84.8	497	358	72.0	358	319	89.1	64.2	71.9
	4703	254	237	93.3	237	221	93.2	87.0	442	335	75.8	335	290	86.6	65.6	73.4
	4704	592	518	87.5	518	484	93.4	81.8	1,005	749	74.5	749	660	88.1	65.7	71.6
	4705	314	292	93.0	292	274	93.8	87.3	505	386	76.4	386	345	89.4	68.3	75.6
	4706	703	543	77.2	543	473	87.1	67.3	1,123	884	78.7	884	772	87.3	68.7	68.2
	4707	281	232	82.6	232	213	91.8	75.8	413	324	78.5	324	287	88.6	69.5	72.0
	4708	249	220	88.4	220	204	92.7	81.9	384	300	78.1	300	270	90.0	70.3	74.9
	4709	327	274	83.8	274	240	87.6	73.4	515	396	76.9	396	358	90.4	69.5	71.0
	4710	299	264	88.3	264	240	90.9	80.3	501	373	74.5	373	333	89.3	66.5	71.6
	4714	.	.	.	.	.	.	.	654	438	67.0	438	383	87.4	58.6	58.6
Alta.	**Total**	6,904	5,621	81.4	5,621	4,978	88.6	72.1	10,553	7,553	71.6	7,553	6,343	84.0	60.1	64.9
	4831	939	813	86.6	813	725	89.2	77.2	1,525	1,094	71.7	1,094	954	87.2	62.6	68.1
	4832	1,595	1,280	80.3	1,280	1,142	89.2	71.6	2,527	1,799	71.2	1,799	1,482	82.4	58.6	63.7
	4833	1,281	1,006	78.5	1,006	886	88.1	69.2	1,961	1,427	72.8	1,427	1,224	85.8	62.4	65.1
	4834	1,647	1,327	80.6	1,327	1,177	88.7	71.5	2,257	1,629	72.2	1,629	1,354	83.1	60.0	64.8
	4835	1,442	1,195	82.9	1,195	1,048	87.7	72.7	2,283	1,604	70.3	1,604	1,329	82.9	58.2	63.8
B.C.	**Total**	8,963	7,150	79.8	7,150	6,364	89.0	71.0	14,369	10,631	74.0	10,631	9,049	85.1	63.0	66.1
	5911	315	276	87.6	276	239	86.6	75.9	525	395	75.2	395	342	86.6	65.1	69.2
	5912	329	282	85.7	282	249	88.3	75.7	528	406	76.9	406	358	88.2	67.8	70.8
	5913	574	512	89.2	512	486	94.9	84.7	1,019	769	75.5	769	677	88.0	66.4	73.0
	5914	537	427	79.5	427	397	93.0	73.9	856	637	74.4	637	541	84.9	63.2	67.3

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
	5921	599	473	79.0	473	408	86.3	68.1	916	716	78.2	716	602	84.1	65.7	66.7
	5922	833	669	80.3	669	606	90.6	72.7	1,394	1,036	74.3	1,036	840	81.1	60.3	64.9
	5923	889	729	82.0	729	649	89.0	73.0	1,462	1,060	72.5	1,060	885	83.5	60.5	65.2
	5931	442	381	86.2	381	353	92.7	79.9	847	579	68.4	579	476	82.2	56.2	64.3
	5932	965	710	73.6	710	654	92.1	67.8	1,580	1,045	66.1	1,045	858	82.1	54.3	59.4
	5933	695	573	82.4	573	466	81.3	67.1	1,020	714	70.0	714	603	84.5	59.1	62.3
	5941	747	588	78.7	588	545	92.7	73.0	1,185	921	77.7	921	807	87.6	68.1	70.0
	5942	546	457	83.7	457	407	89.1	74.5	854	676	79.2	676	595	88.0	69.7	71.6
	5943	283	188	66.4	188	182	96.8	64.3	440	346	78.6	346	302	87.3	68.6	66.9
	5951	374	291	77.8	291	247	84.9	66.0	515	369	71.7	369	317	85.9	61.6	63.4
	5952	510	362	71.0	362	295	81.5	57.8	736	576	78.3	576	519	90.1	70.5	65.3
	5953	325	232	71.4	232	181	78.0	55.7	492	386	78.5	386	327	84.7	66.5	62.2
Y.T.	**6001**	1,234	1,023	82.9	1,023	930	90.9	75.4	327	263	80.4	263	240	91.3	73.4	75.0
N.W.T.	**6101**	1,402	1,200	85.6	1,200	1,027	85.6	73.3	234	185	79.1	185	161	87.0	68.8	72.6
Nvt.	**6201**	938	807	86.0	807	725	89.8	77.3	.	.	.	.	.	.	.	77.3

Appendix G (2011-2012) - Table 9.4 Response rate by Local Health Integrated Network (LHIN) and frame in Ontario

Geography		Area frame							Phone frames							Combined
Prov. Terr.	Health Region	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	# in scope HH	# resp. HH	HH resp. rates	# pers. select.	# resp.	Pers. resp. rates	Resp. rates	Combined resp. rates
Ont.	**Total**	24,559	19,559	79.6	19,559	17,367	88.8	70.7	38,865	30,035	77.3	30,035	25,548	85.1	65.7	67.7
	3501	1,605	1,324	82.5	1,324	1,218	92.0	75.9	2,702	2,094	77.5	2,094	1,783	85.1	66.0	69.7
	3502	2,535	2,140	84.4	2,140	1,995	93.2	78.7	4,418	3,521	79.7	3,521	3,040	86.3	68.8	72.4
	3503	1,374	1,125	81.9	1,125	1,006	89.4	73.2	2,111	1,688	80.0	1,688	1,430	84.7	67.7	69.9
	3504	2,976	2,323	78.1	2,323	2,023	87.1	68.0	4,569	3,526	77.2	3,526	3,001	85.1	65.7	66.6
	3505	1,225	1,004	82.0	1,004	875	87.2	71.4	1,946	1,412	72.6	1,412	1,132	80.2	58.2	63.3
	3506	1,282	1,086	84.7	1,086	988	91.0	77.1	2,073	1,560	75.3	1,560	1,270	81.4	61.3	67.3
	3507	1,290	965	74.8	965	847	87.8	65.7	2,072	1,378	66.5	1,378	1,109	80.5	53.5	58.2
	3508	1,697	1,350	79.6	1,350	1,174	87.0	69.2	2,643	1,814	68.6	1,814	1,418	78.2	53.7	59.7
	3509	2,337	1,825	78.1	1,825	1,605	87.9	68.7	3,710	2,828	76.2	2,828	2,381	84.2	64.2	65.9
	3510	1,462	1,181	80.8	1,181	1,003	84.9	68.6	2,213	1,790	80.9	1,790	1,588	88.7	71.8	70.5
	3511	2,396	1,905	79.5	1,905	1,681	88.2	70.2	3,465	2,875	83.0	2,875	2,511	87.3	72.5	71.5
	3512	1,154	836	72.4	836	741	88.6	64.2	1,834	1,463	79.8	1,463	1,282	87.6	69.9	67.7
	3513	2,017	1,575	78.1	1,575	1,412	89.7	70.0	3,311	2,663	80.4	2,663	2,336	87.7	70.6	70.3
	3514	1,209	920	76.1	920	799	86.8	66.1	1,798	1,423	79.1	1,423	1,267	89.0	70.5	68.7