Canadian Longitudinal Study on Aging
Étude longitudinale canadienne sur le vieillissement

CLSA TECHNICAL DOCUMENT 2025

Sampling and Computation of Response Rates and Sample Weights for the Tracking (Telephone Interview) Participants and Comprehensive Participants

Mary Thompson, Changbao Wu, Urun Erbas Oz, Lauren E. Griffith, Christina Wolfson, Parminder Raina

v4.0 2025 May 01

Contents

1. BACKGROUND

The Canadian Longitudinal Study on Aging (CLSA) recruited 51,338 Canadian residents aged 45-85 years at baseline to be followed for at least 20 years or until death or loss to follow-up. There are two components to the sample: (1) CLSA Tracking, whose target was 20,000 people from across the 10 Canadian provinces; and (2) CLSA Comprehensive, which aimed to recruit 30,000 people living within 25-50 km of one of 11 Data Collection Sites (DCS) across 7 Canadian provinces. Since the Canadian Community Health Survey (CCHS) on Healthy Aging was the first source of participants for CLSA Tracking, the CLSA used the same selection criteria for the recruitment of all participants as the CCHS Healthy Aging Cycle 4.2 survey. The CCHS-Healthy Aging (CCHS-HA) sample is a nationally representative sample of people over the age of 45. Excluded from the sampling frame, and consequently the CLSA, are residents in the three territories, persons living on federal First Nations reserves and other First Nations settlements in the provinces, full-time members of the Canadian Armed Forces, and individuals living in institutions. This latter exclusion means that individuals living in long-term care institutions (i.e., those providing 24-hour nursing care) are excluded from the CLSA at baseline. Individuals living in households and transitional housing arrangements (e.g., seniors' residences, in which only minimal care is provided) are included at baseline. Individuals with cognitive impairment, as determined by CLSA interviewers, and those unable to respond in English or French are also excluded from the CLSA. The CLSA team also used three additional sampling sources to select a random sample for both CLSA Tracking and Comprehensive cohorts. The additional sampling frames were Provincial Health Registries (HR), Telephone sampling (TS), and the Quebec Longitudinal Study on Nutrition and Aging (NuAge). The inclusion criteria used to select participants from these three sampling frames were the same as for CCHS-HA.

2. RESPONSE RATES

We defined response rates as the number of participants divided by the estimated number of those sampled who were eligible.

Detailed information about the response rate calculations is given in section 6.

3. SAMPLE WEIGHTS

It is standard practice in surveys to use sampling weights. Each participant in the study is assigned a sample inflation weight constructed based on the inclusion probability. The inflation weights provided with the data aim to provide researchers with an estimate of how many people in each province (and in Canada) are represented by each CLSA participant. The aim is to ensure that, when estimating the mean value of some variable or the proportion with some characteristic, the value obtained is representative of the eligible provincial (and Canadian) population.

The weighting is necessary because the probability of selecting individuals from certain sub-groups of the population varied. As well, the probability that those selected agreed to participate varied within groups. For example, the CLSA ultimately included a much higher proportion of people in PEI than of people in Ontario. If the sample inflation weights are not used, any estimate of Canadian population means or proportions will be skewed toward the mean proportion for PEI. Using the weights in the calculations will remove this bias.

3.1. Inflation Weights

The CLSA Tracking cohort and Comprehensive cohort inflation weights were constructed to account for sample misrepresentation resulting from unequal sampling probabilities, frame coverage error and non-response, and to improve the precision of estimates through the use of auxiliary information.

First, the basic design weights were computed proportional to the reciprocals of the individual inclusion probabilities; they were then re-calibrated to the sum of the targeted (eligible) Canadian population using benchmarks based on Statistics Canada data.

Inflation weights were also provided for the pooled sample from the two cohorts.

From all of these calculations, and for each of the Tracking or Comprehensive or pooled samples, the first wave inflation weight w_i of unit i [WGHTS_INFLATION_TRM, WGHTS_INFLATION_COM, WGHTS_INFLATION_CLSAM] is interpreted as the number of persons in the population that unit represents, and the sum of the w_i over all i in the sample equals the known or assumed population size.

3.2. Analytic Weights

For analytical purposes, some different considerations apply. Regression and logistic regression analyses may be designed to estimate the relationships among variables, not so much for the population at hand as for a hypothetical population of like people, represented by the sample.

The so-called analytic weights supplied with the Tracking cohort data [WGHTS_ANALYTIC_TRM] are proportional to the inflation weights but rescaled to sum to the sample size within each province, so that their mean value is 1 within each province.

The analytic weights supplied with the Comprehensive cohort data [WGHTS_ANALYTIC_COM] are proportional to the inflation weights but rescaled to sum to the sample size within the individual DCS part of each province, so that their mean value is 1 within each individual DCS.

The analytic weights were also provided for the pooled data [WGHTS_ANALYTIC_CLSAM] with the mean value of 1 within each individual DCS and provincial Non-DCS area.

3.3. When and How to Use the Weights

For the estimation of a descriptive parameter of the finite study population, the inflation weights should be used. For analyses that examine relationships between variables at the national or provincial level, analytic (rescaled) weights should be used. For analyses of smaller sub-groups, the analytic weights are likely to be appropriate. However, consultation with a statistician is recommended.

The detailed information about the sample weights calculations is in section 7.

3.4. Longitudinal Weights

The longitudinal baseline-to-FUP1 weights are to be used when a FUP1 outcome is being modelled or predicted in terms of baseline explanatory variables and/or baseline covariates; in such cases the analysis is restricted to individuals with information available at both waves. Examples include (a) modeling death by FUP1 in terms of baseline predictor variables; (b) modeling hypertension at FUP1 in terms of baseline explanatory variables, given survival to FUP1.

4. PRIMARY SAMPLING UNIT and SAMPLING STRATA VARIABLES

The use of complex survey software, as available for example in SAS, SPSS, Stata and R, is recommended for analyses, so that the sampling design can be accounted for. This will require specification of the appropriate weights variable (given above), and of characteristics of the sampling design, namely strata and primary sampling units or PSUs.

A stratified sampling design involves dividing the population into mutually exclusive (non-overlapping) strata, and a sample is taken from every stratum. Within strata, individuals may be selected directly (single stage sampling) with a probability design. Alternatively, the sampling may be done in multiple stages within geographic strata.

If the sampling is done in stages, the PSU is a geographic unit selected by probability sampling at the first stage of sampling, within geographic strata. Within each selected PSU, there is a (possibly multi-stage) design for sampling individuals. This sample structure means that selected individuals are “clustered” into PSUs. Because individuals who live closer together tend to be more alike, a design with this kind of clustering tends to lead to less efficient estimation than does single stage probability sampling.

The samples from the HR and the TS frame are effectively single stage, with no geographic clustering of respondents. The CCHS-HA design had at least two stages within health regions, but because we have not been provided with PSU information, and because the sample size is small enough that individuals in the same age group are not likely to be clustered together much more than in a single stage design, we take the CCHS-HA design also to be single stage. For the purposes of specification of the design in complex survey software, the PSU should be taken to be the individual, as represented by the unique ID variable, in the CLSA data for each cohort and for the pooled data.

For the strata variable to be specified in complex survey software, we recommend using the geographic strata variables, namely GEOSTRAT_TRM (10 provinces crossed with DCS/Non-DCS such as AB_DCS, ON_DCS, or BC_Non-DCS), GEOSTRAT_COM (7 provinces crossed with individual DCS such as AB_Calgary), GEOSTRAT_CLSAM (10 provinces crossed with individual DCS and provincial Non-DCS such as AB_Calgary or AB_Non-DCS) for the Tracking, Comprehensive and pooled data, respectively. This is essentially the same geographic stratification for all three data sets.

Detailed information about sampling and strata determination is given in section 5.

5. SAMPLE

The sample was obtained via four sources:

1. Canadian Community Health Survey – Healthy Aging (CCHS): only for CLSA Tracking cohort
2. Provincial Health Registries (HR)
3. Telephone Sampling (TS)

4. Quebec Longitudinal Study on Nutrition and Aging (NuAge): only for CLSA Comprehensive cohort, in Quebec

For each we obtained 'pre-recruits', that is, people expressing preliminary interest in participation. Pre-recruits provided contact information. CLSA staff at Computer Assisted Telephone Interview (CATI) sites[1] attempted to talk to each person by phone, explain the CLSA in more detail, and describe what participation would entail, and asked if the person would participate. Those who agreed to do so were considered 'recruits'.

Tracking Cohort: The full interview ('60 minute interview') was conducted immediately or later. Once a person completed the interview, s/he was called a 'provisional participant'. Consent forms could have been received by the CLSA before the 60 minute interview, but usually were not sent back until after the interview. People were not considered to be 'participants' until both the interview was completed and written consent was received by the CLSA.

Comprehensive Cohort: A person who completed the in-home interview and consent form (consent form was signed when the interview questionnaire was administered), but had not yet completed the DCS assessment was called a 'partial participant'. People were not considered to be 'participants' until the DCS assessment was completed.

Section 5.2. describes how people were recruited in detail for each source.

5.1. Criteria for the Sample

Tracking Cohort: The total of 20,000 participants was to be divided among the provinces to allow reasonably precise estimates of various parameters or associations to be made at the provincial level, while obtaining more of the sample from the larger provinces. The minimum target sample in any province was 1,100 in Prince Edward Island, and the maximum was 4,388 in Ontario. As well, the sample was distributed by age and sex within provinces. Eight strata were formed, based on age (45-54, 55-64, 65-74, and 75-85) and sex (male or female). Across the country, the four younger age-sex strata were to include 3,000 each and the four older age-sex strata 2,000 each.

Comprehensive Cohort: The total of 30,000 participants was to be divided among the provinces to allow reasonably precise estimates of various parameters or associations to be made at the provincial level, while obtaining more of the sample from the larger provinces. The minimum target sample in any province was 3,000 in Alberta, Manitoba, Newfoundland and Labrador, and the maximum was 6,000 in British Columbia, Ontario, and Quebec. Eight strata were formed, based on age (45-54, 55-64, 65-74, and 75-85) and sex (male or female).

Table 1 and Table 2 show the target and actual numbers of participants in each stratum, respectively, for the CLSA Tracking and Comprehensive cohorts.

5.2. Sources for the Sample (Sampling Frames)

5.2.1. Canadian Community Health Survey – Healthy Aging (CCHS): Only for Tracking Cohort

[1] The CATI sites, operated by the CLSA, conducted the 60-minute telephone interviews.

The first participants in the CLSA were recruited from participants in the CCHS-HA which was conducted in 2008-2009 by Statistics Canada. As described above, the inclusion and exclusion criteria for the CLSA were adapted from those used in the CCHS. Participants in the CCHS aged 45-85 were asked if they would provide consent to allow Statistics Canada to pass on their names and contact details to the CLSA; they were also asked to give permission for passing on the data they provided as part of the CCHS interview to the CLSA. The disposition of participants in the CCHS who provided contact information and/or the data to the CLSA team is given in Table 3. Those who provided their contact information were approached by the CLSA. They were first sent an information package describing the purpose of the study, the criteria for participation, and a consent form by mail. They were called up to 10 times[2] to be invited to join the study. Those who completed the CLSA baseline questionnaire[3] and provided written consent were considered CLSA Tracking participants.

5.2.2. Provincial Health Registries (HR)

In the second approach, the provincial government departments or data stewards responsible for housing the healthcare administration databases mailed the information packages directly to the randomly chosen age-eligible persons on behalf of the CLSA.

Depending on provincial requirements, either the introductory letter included in the information package was signed jointly by a provincial government representative designated by the province in question and the lead principal investigator (PI) for the CLSA, or separate introductory letters from the CLSA and/or Ministry were included. The letter package contained the introductory letter, a brief explanation of the CLSA, a consent form to be contacted, and a stamped, addressed envelope for potential participants to mail back to the CLSA. A reminder letter was sent twenty days after the initial mail-out. Those who replied and agreed to be contacted were called by the CLSA team. They were given further information and were asked to join the study.

For the Tracking cohort, eight provinces – British Columbia (BC), Manitoba (MB), New Brunswick (NB), Newfoundland and Labrador (NL), Nova Scotia (NS), Ontario (ON), Prince Edward Island (PE) and Saskatchewan (SK) – agreed to use their registry to select a sample and send letters to potential participants.

For the Comprehensive cohort, five provinces – BC[4], MB, NL, NS, and ON – agreed to use their registry to select a sample and send letters to potential participants.

There were two mailings. The first was sent to a stratified random sample of all eligible people in the province.[5] The second was sent to a more targeted sample as described below under ‘Targeted Samples’.

[2] In practice, for some people, more than 10 attempts were made to contact them by phone.

[3] Because the CLSA received the contact information for CCHS participants prior to the launch of the CLSA recruitment, CCHS-based participants first completed a 20-minute pre-recruitment questionnaire prior to completing the 60-minute baseline questionnaire.

[4] For the CLSA comprehensive, only the males aged between 45 and 54 were targeted in BC.

[5] This assumed each registry contained a complete list of people in the province. This will have excluded some eligible people who were not registered.

For the Tracking cohort, SK and ON sent the first mailing but not the targeted one, while BC sent the targeted one but not the first.[6]

For the Comprehensive cohort, only NS sent both first mailing and the targeted one, while BC, MB, NL, and ON sent only the first.

The CLSA team provided the number of people to be sent a letter in each age and sex stratum to the provinces; numbers were calculated taking into account predicted response rates, based on reported response rates of other studies in Canada, and pilot work for the CLSA. The ministries randomly selected people in each stratum, where possible taking account of the exclusion criteria. If more than one individual from a household was chosen, some provinces (MB, NL, SK, ON, PE, and BC) were able to select one person randomly to be sent the CLSA package while the other two provinces (NB and NS) were not able to do so.

The procedure used for sample selection varied by individual province. Although the same exclusion criteria applied to each province, the exact sampling process was based on the province's database/system.[7]

The dates of mail-outs and the numbers of people who were sent letters by province, age and sex are given in Tables 5, 6, 7 and 8.

5.2.3. Telephone Sampling-Random Digit Dialing (RDD)

Random digit dialing is a procedure in which valid telephone numbers[8] are generated randomly to draw a sample of households, which are then called by telephone. The CLSA used RDD as a third sampling approach in all provinces for the Tracking cohort, except NS where the target number of participants was achieved through mail-outs alone; and RDD was used as a second sampling approach in seven provinces which have DCSs for the Comprehensive cohort. Table 9 shows the data collection sites by province. Area codes, specific to provinces except for the code 902 which includes both NS and PE, were used to sample telephone numbers from each province.

Given various difficulties in calling cell phones, only landlines were included in the study. Excluding cell phones could have created an important bias if the study would not reach a large proportion of the eligible population. However, a survey of residential telephone service conducted by Statistics Canada in 2010 showed that very few households with members over the age of 45 did not have a landline. Based on these data we estimated that roughly 5% of potential participants would be excluded from sampling using RDD, and we considered this to be acceptably low.

[6] In BC the process of pre-recruitment was slightly different in that after completing the sample selection, the Ministry provided the contact information to BC-based CLSA researchers who sent invitation letters directly.

[7] For example, NL did not exclude people who lived in long-term care facilities while the other provinces could do this. Of the provinces that did mail-outs, only NB and PE provided the number of packages that were returned as undeliverable, presumably because the selected people had moved or died. More details on the exclusion criteria that the provinces were able to implement are given in Table 4.

[8] Some numbers are known not to be valid, e.g., area codes and exchanges cannot begin with the numerals 0 or 1 and can be excluded from the list of possible numbers to call.

After a pilot study had shown that RDD was feasible, a well-known Canadian professional polling company, Leger, was hired to conduct the RDD. A script was developed by the CLSA team for the company to use during the initial telephone interview to select and pre-recruit people. The CLSA team provided the company with quotas for the number of people in each stratum to be pre-recruited. The numbers were based on the target sample sizes remaining in those provinces, and on the assumption that 40% of Tracking pre-recruits would later agree to participate in the Tracking cohort and 60% of Comprehensive pre-recruits would later agree to participate in the Comprehensive cohort. The pre-recruitment was spread out over time, via weekly quotas for the polling company, to ensure that CATI sites did not have a backlog of people to call and thus ensure a relatively short time gap between pre-recruitment by RDD and further contacts with the CLSA. Telephone numbers that were not businesses and appeared to be valid numbers were called up to 10 times before being considered to be non-respondents for pre-recruitment.

During the interview, information on whether there was anyone in the household aged between 45 and 85 and a roster of people in that age range were obtained. One person was randomly selected from the roster as a potential participant and the other eligibility questions were asked of the potential participant. Household members whose age-sex quota had already been filled were excluded from selection. Those who agreed to be contacted by the CLSA team were termed 'pre-recruits'.

5.2.4. Targeted Samples

Early analyses on initial recruits showed under-representation of people with lower levels of education, a marker for various risk factors (many of which are assessed in the CLSA). This under-representation could potentially lead to low statistical power to identify relationships between these variables and health outcomes. Thus, to increase heterogeneity in the independent variables in the Tracking cohort, the CLSA chose to over-sample people from dissemination areas[9] (DAs) with relatively high proportions of people with lower levels of education. We anticipated that a similar problem would also arise later in sampling for the CLSA Comprehensive participants. To ensure we would still have enough participants in the catchment area around the DCS when sampling for the CLSA Comprehensive, we chose to conduct the targeted sampling for the CLSA Tracking in areas outside the catchment area.[10] More detail on the determination of these DAs is in Appendix 1. We treated the extra mail-outs from HRs as additional, but different samples from the same sampling frame. We also had random (telephone) sampling from listed telephone numbers (RTS) – which was done by CLSA CATI sites – to increase the proportion of people with lower levels of education as a different sample. To ensure the people called lived in one of the identified DAs, listed telephone numbers, which included addresses, were used as the sampling frame. For the Comprehensive cohort, although there were extra mail-outs from HRs and sampling from listed telephone numbers, there was no oversampling from lower education DAs.

[9] Statistics Canada states "[a] dissemination area (DA) is a small, relatively stable geographic unit composed of one or more adjacent dissemination blocks. It is the smallest standard geographic area for which all census data are disseminated." Each area covers between 400 and 700 people, and this allowed us to identify areas with relatively high proportions of people with lower levels of education.

[10] Except for NS where recruitment had already been done before the decision to oversample exclusively from Non-DCS areas.

The oversampling in the Tracking cohort meant that the target of 50,000 people was exceeded. In total, there were 21,241 participants in the CLSA Tracking cohort, and 30,097 participants in the CLSA Comprehensive cohort.

5.2.5. Quebec Longitudinal Study on Nutrition and Aging (NuAge)[11]: Only for Comprehensive Cohort

The last participants in the CLSA were recruited from the NuAge study, which is a longitudinal study on nutrition as a determinant of successful aging and includes a cohort of 900 healthy men and 900 healthy women born between 1921 and 1935 to be monitored annually for a period of 5 years. Participants in NuAge were asked by the NuAge investigators if they would provide consent to share their information with the CLSA. Only the participants falling into the age group 75-85 who provided their contact information were approached by the CLSA. They were first sent an information package describing the purpose of the study and the criteria for participation by mail.

5.3. Sample Components in the CLSA

These approaches resulted in the CLSA samples having several components, based on sample source:

- CCHS: only for Tracking cohort
- HR
  - HR1 – initial Health Registry mail-outs
  - HR2 – targeted Health Registry mail-outs
- Telephone Sampling (TS)
  - RDD – done by Leger
  - RTS – targeted, conducted by CLSA CATI
- NuAge: only for Comprehensive cohort

Tables 10 and 11 give the number of participants in each cohort crossed by sample source and province.

6. CALCULATION of RESPONSE RATES

For each of the cohorts, response rates were computed for each sample component separately, and overall.

CCHS: Statistics Canada provided the CLSA with response rates by province and age group for CCHS participants identified through the CCHS. The rates were not broken down by sex. We calculated the 'contact sharing rate' as the proportion of people in each age and sex group who gave permission to Statistics Canada to pass on their contact information to the CLSA. The CLSA computed recruitment and participation rates among those people who permitted Statistics Canada to pass their contact information to the CLSA. The rates were based on whether people completed the 20-minute interview (recruits) and then completed the 60-minute interview and provided consent forms (participants), respectively. Ineligible people were removed from denominators when calculating the rates; for those who could not be reached/contacted, estimates of the numbers ineligible were used, assuming the proportion of ineligibles in those not contacted was the same as in those who were contacted. Recruitment and participation rates were computed by province, age and sex. Overall response rates for each

[11] For the details, please go to http://www.rqrv.com/en/init_NuAge.php.

province, sex, and age group were obtained by multiplying four rates: the CCHS response rate, the contact sharing rate, the recruitment rate and the participation rate. The details of the calculations are shown in Appendix 2.

HR: For the health registries, we used the numbers of people who were sampled and were sent letters, replied to the CLSA, were contacted by the CLSA and agreed to join the CLSA in each province, age and sex group. We also had the number of mailed packages returned as undeliverable in some provinces. As well, after contacting some people we discovered they did not meet the CLSA eligibility criteria. We computed the proportion of those sent letters who returned the contact form, and the proportion of those who completed the 60-minute interview and returned their completed consent form for Tracking, or completed the in-home interview, consent form, and DCS assessment for the Comprehensive cohort. Response rates were obtained by multiplying proportions for each province, age and sex group. Calculations adjusted denominators to account for the numbers of returned packages and ineligible people. The formulas are shown in Appendix 3.

There were some people selected by one province's health registry who had moved to a different province. For confidentiality reasons, the CLSA did not have a list of people selected for the mail-outs in each province and could not necessarily distinguish people who had moved to a different province. Such people, likely a very small proportion of the total,[12] were allocated to the province to which they had moved, rather than the one from which they had been selected. This was done for those who moved to a province that had also conducted a mail-out. For those who moved to a province that had not conducted a mail-out, it was inferred that they had moved from another province. To compute response rates and sample weights, we used data from Statistics Canada on interprovincial migrants (see footnote 12) and weighted random numbers to assign them to a province that had conducted a mail-out.

Telephone Sampling: In RDD, phone numbers were called up to 10 times and each call result was coded by the interviewer (answering machine, line busy, no answer, no one in that age group, refusal, language barrier, etc.). The company that conducted the RDD provided two files to the CLSA for each province: a call history file including the codes for each call attempt and a file including the results of initial interviews (including answers to the eligibility questions). A new file for each province was created by the CLSA team merging these two files; after this, call dates were sorted within phone numbers and the last call was identified for each phone number (household). From these last calls, the numbers of valid phone numbers, answered phone numbers, residential phone numbers, households for which age-eligibility was established, and age-eligible households were obtained. The numbers of rostered households and rostered people (in rostered households) were determined. The file also provided the numbers of selected people, people who agreed to participate in the study, people who were eligible, and people who provided contact information to the CLSA team. These numbers were used to calculate response rates for each province, sex, and age group. Since we did not have age and sex breakdown for household enumeration rate, the same household enumeration rate was used for all age and sex groups within a province.

[12] The total number of inter-provincial migrants in 2010/11 was roughly 257,000, less than 1% of the Canadian population (http://www.statcan.gc.ca/pub/91-209-x/2013001/article/11786/tbl/tbl3-eng.htm) and concentrated in those younger than the age-eligible range for the CLSA (http://www.statcan.gc.ca/pub/91-209-x/2013001/article/11786/fig/fig2-eng.htm).

RTS was conducted by the CLSA CATI sites. The process was a little different. Once a number had been called, the system did not call that number again until all the sampled numbers had been called. As a result, some phone numbers were called only once. In addition, the codes used for call results were different from the ones in RDD. However, the same procedure was used to compute the response rates.

NuAge pre-recruits were treated as RDD pre-recruits and counted in RDD response rates.

The details of the calculations for telephone sampling response rates are shown in Appendix 4.

Pooled response rates were also calculated by province for each sampling frame (not calculated for CCHS as it is only in Tracking cohort) and overall, repeating the same calculation steps for the pooled numbers.

Response rates are shown in Tables 12, 13, and 14 for Tracking cohort, Comprehensive cohort, and pooled cohort, respectively.

7. CALCULATION of BASELINE SAMPLE (INFLATION) WEIGHTS

Original Baseline Sample Weights: The CCHS Healthy Aging (CCHS-HA) 2008-2009 survey represents the Canadian population aged 45-85 years in 10 provinces. CCHS-HA participants were asked if they would agree to share their contact information with the CLSA and were used as the first sampling frame for the tracking participants in the CLSA. As a result, the CLSA recruitment criteria were based on those of the CCHS-HA. For CLSA tracking participants who were recruited from CCHS-HA, we used their CCHS-HA sampling weight as a basis for CLSA sampling weight. Because the CCHS-HA represented the same target population as the CLSA, we also used the CCHS-HA weights to calibrate the CLSA tracking cohort weights using benchmarks for CCHS strata. We then applied the same strategy for the comprehensive cohort which includes CLSA participants sampled from a geographic catchment area around one of eleven data collection sites located in seven provinces. It was understood that the CCHS-HA weights were generated in such a way that the sample would represent each health region. The CLSA comprehensive catchment area were generally an area within a 25 km radius of the DCS (in some cases up to 50 km) which often covers the main health region and portions of neighbouring health regions, making it difficult to calibrate CLSA sampling weights properly for the comprehensive cohort. In some sites, the catchment area was quite different from a union of health regions. One project evaluated how weighted estimates of socio-demographic characteristics from the CLSA comprehensive sample compared with those from 2016 census data for one of the cities with a DCS. While the tracking and combined sample are fairly representative of the target population at the province level, this project identified that the weighted comprehensive sample was less representative of the target population with respect to sex and age group at the city-level. We also recognized that the CLSA, despite efforts to over-sample in areas with higher proportions of people with lower levels of education (< high school), had lower levels of endorsement of low income or education, meaning that without taking this under-representation into account, even weighted estimates would be biased for prevalence estimation and/or estimates of associations likely to be influenced by income or education.[1] For these reasons we decided to recalculate our sampling weights.

New Baseline Sample Weights: CLSA participants were recruited between 2011 and 2015, making 2011 census data most suitable for calibrating our baseline weight. Because we had less missing information on education than income, we chose to use individual-level self-reported education (by geography) in addition to sex and age group (by geography), as a calibration variable. In constructing the calibration benchmarks, the best approach for geography would be to use postal code level information because the DCS catchment areas was defined using distance from the DCS determined by postal codes. Ideally, we would have been able to use census data for all the benchmarks. However, in 2011 only the short form was used in the census, and individual education information was not included there. The most suitable alternative to census data were data from the National Household Survey (NHS) 2011, the replacement for the long form census, which was distributed to about 35% of the census population. However, the 2011 NHS data do not provide postal code level information, and the lowest level geographic unit was the census dissemination area. Therefore, in our new weight calculation, we used dissemination-area-level aggregated information for the individual level age-sex and education benchmark calculations.

For the detailed information on the old CLSA sample weights, please go to https://www.clsa-elcv.ca/doc/1041.

7.1. Source of Benchmarks for Calibration of Design Weights: National Household Survey 2011

The sample is designed to represent the population of interest, but it is almost never fully representative even within sampling strata (geographic strata crossed with sex crossed with age group). Sample weights are used to make statistics computed from the data more representative of the population. We used the National Household Survey (NHS) 2011 dataset, a replacement for the 2011 long form census, as our population reference for calibration purposes.

Census 2011 was a mandatory short survey including every (eligible) person in Canada while the NHS 2011 was a voluntary long survey (including the questions from the short form survey) and distributed to about 30% of Canadian census households (response rate of 68.6%.)

"The NHS covers all persons who usually live in Canada, in the provinces and the territories. It includes persons who live on Indian reserves and in other Indian settlements, permanent residents, non-permanent residents such as refugee claimants, holders of work or study permits, and members of their families living with them.

Foreign residents such as representatives of a foreign government assigned to an embassy, high commission or other diplomatic mission in Canada, members of the armed forces of another country stationed in Canada, and residents of another country who are visiting Canada temporarily are not covered by the NHS.

The survey also excludes persons living in institutional collective dwellings such as hospitals, nursing homes and penitentiaries; Canadian citizens living in other countries; and full-time members of the Canadian Forces stationed outside Canada. Also excluded are persons living in non-institutional collective dwellings such as work camps, hotels and motels, and student residences.

A survey's reference date is the date to which respondents refer when answering the questions. The reference date of the NHS is May 10, 2011, the date of the 2011 Census of Population."[13]

After excluding households on First Nations reserves and Aboriginal Inuit areas of residence, people who were not permanent residents, people who lived in the 3 territories, and people younger than 45 or older than 85

[13] National Household Survey User Guide: Survey content and target population https://www12.statcan.gc.ca/nhs-enm/2011/ref/nhs-enm_guide/guide_1-eng.cfm, 09.03.2020

from NHS 2011 dataset, we ended up with 2,711,183 people with a total weight of 13,655,038.79, summing the NHS 2011 inflation weights.

7.2. Calculation of Initial Design Weights

7.2.1. Tracking Cohort

CCHS: Statistics Canada computed the weights for those who allowed Statistics Canada to pass on survey and contact information ('survey+contact' group) or survey data only ('survey only' group) to the CLSA; the weights added up to 13,232,650.77[14] (Table 15).

The CLSA team attempted to contact all 12,269 CCHS participants who provided their contact information. This included 11,742 'survey+contact' CCHS participants and 527 other CCHS participants who allowed Statistics Canada to pass on their contact information, but not their survey data ('contact only' participants). For the contact only participants, the CLSA imputed sampling weights. First, sampling weights for the 'survey+contact' and 'survey only' CCHS participants were grouped by the sampling strata mentioned above. The imputed weight for a 'contact only' participant was the median[15] weight of the relevant stratum. The numbers of participants within the sampling strata are given in Tables 16 and 17.

Statistics Canada supplied weights were highly variable. We created a table of CCHS basic weights by province by sex crossed with age decades. If the "maximum/median" ratio of the stratum (province, age, sex) was higher than 3, then we set the weights above 3×median weight in the respondent's stratum to be 3×median weight.

HR: We used the NHS weights by province and sex crossed with age decades (45-54, 55-64, 65-74, and 75-85) as well as the LowEd/Non-LowEd breakdown (used in determining where to take the targeted samples) to build the basic design weights.

We considered the "stratum" h to be sex crossed with age decade within province:

In stratum h, let N_hL be the NHS 2011 number in stratum h , LowEd area, and N_hnL be the NHS 2011 number in stratum h, Non-LowEd area. Suppose the sum of these two is N_h.

Let n_0h be the initial sample size in stratum h, and let n_hL be the later "targeted" sample size.

If individual i in the stratum h is not in the LowEd area, the basic weight is the reciprocal of the inclusion probability, namely (N_h/n_0h). If individual i in stratum h is in the LowEd area, the basic weight is again the reciprocal of the inclusion probability, and the inclusion probability is approximately $(n\_0h/N\_h) + (n\_hL/N\_hL) - (n\_0h/N\_h) \times (n\_hL/N\_hL)$.

In this formulation, it is assumed that all the members of stratum h are eligible and that non-response has occurred at random within strata. A refinement would have been to replace N_hL and N_hnL by estimated numbers eligible in LowEd and Non-LowEd parts, but this was not done.

[14] For further information, http://www12.statcan.gc.ca/census-recensement/index-eng.cfm, 04.11.2014.
[15] Due to the presence of extreme weights within stratum, median weight was prefered to use instead of mean weight.

Example: Initial weights for female 45-54 Manitoba HR mail-out participants

NHS	CLSA first mail-out participants	CLSA targeted mail-out participants
ALL, N_h	n_0h	n_hL
85,740	77	
LowEd, N_hL	LowEd, n_0h	
5,650	28	4
Non-LowEd, N_hnL	Non-LowEd, n_0h	
80,090	49	

Initial weight for the first mail-out participants who were NOT in the LowEd areas: $1/(77/85{,}740) = 1{,}113.5$.

Initial weight for the first mail-out participants who were from the LowEd areas and for the targeted mail-out participants: $1/\big((77/85{,}740) + (4/5{,}650) - (77/85{,}740) \times (4/5{,}650)\big) = 622.9$.

It should be noted that the CLSA did not conduct the first HR mail-out in BC (no participant from the first mail-out), so for the purpose of initial weight assignment, BC targeted HR mail-out participants were considered as Telephone Sampling (TS) participants (LowEd RDD) and the basic weight 1 was assigned to them. In ON and SK, the CLSA conducted the first HR mail-out but not the targeted one; we used the same basic weights for all HR participants in a given age-sex group. Also, because there was no targeted mail-out in "NB male 55-64" and "NL female 45-54", we used the same initial weights for the first mail-out participants in these strata regardless of whether or not they were in LowEd areas.

Telephone Sampling: The totals of NHS weights by province crossed with LowEd/Non-LowEd breakdown were used to build the basic design weights.

In telephone sampling, one age-eligible person per co-operating household was selected randomly to participate in the study. Within a contacted and co-operating household, the probability of selecting an individual is $1/k$ where k is the number of age-eligible people living in the same household. We checked the distribution of k by province for both RDD and RTS; we capped k at 3 if $k > 3$. Thus, the *within-household weight* for an individual is taken to be k, which is the reciprocal of the within-household inclusion probability; it is the number of age-eligible people in the respondent's contacted and co-operating household, obtained by the interviewers during the initial telephone interview performed to select and pre-recruit people, and capped at 3.

We needed to multiply this individual within-household weight by something that is proportional to the household weight:

The effective probability of inclusion of a household in the RDD non-targeted sample is approximately the number n_T0 of households from which an RDD respondent came in the province, divided by the number N_T0 of landline telephone households with eligible members in the province (N_T0 is unknown). So, for a non-LowEd area household with eligible members, its effective (rough) probability of inclusion is n_T0/N_T0.

For a LowEd household with eligible members, its effective (rough) probability of inclusion is $(n\_T0/N\_T0) + (n\_TL/N\_TL) - (n\_T0/N\_T0) \times (n\_TL/N\_TL)$, where n_TL is the number of supplementary sample LowEd area households from which a participant came in the province and N_TL is the number of landline telephone numbers in the LowEd area (N_TL is also unknown).

If we ignore the last term, which is likely to be close to negligible, then for a LowEd area household with eligible members, its effective (rough) probability of inclusion is $(n\_T0/N\_T0) \times (1 + A)$ where $A = (n\_TL \times N\_T0)/(n\_T0 \times N\_TL)$.

We needed initial weights only up to a constant of proportionality. Thus, we aimed to estimate A, then take the initial weight for a respondent in the Non-LowEd area to be the individual weight k, and take the initial weight for a respondent in the LowEd area to be $k/(1 + A)$.

We know n_TL, n_T0, but not the ratio N_T0/N_TL. The ratio of the number of eligible households in the province (regardless of telephone status) and the number of eligible households in the LowEd areas (regardless of telephone status), would be a good estimate of N_T0/N_TL, but these numbers were not available. Thus, we estimated N_T0/N_TL more crudely by M_0/M_L, where M_0 is the number of eligible people in the province and M_L is the number of eligible people in the LowEd area, according to the 2011 census.

Example: Initial weights for TS participants in AB

NHS	RDD participants	RTS participants
N_h		
M_0 (overall)	n_T0	n_TL
1,285,535	1,688	67
N_hL		
M_L (overall)		
93,085		

$A = (n\_TL \times M\_0)/(n\_T0 \times M\_L) = (67 \times 1{,}285{,}535)/(1{,}688 \times 93{,}085) = 0.55$

$1 + A = 1.55$

Initial weight for the RDD participants who were not from the LowEd areas=k, number of age eligible people living in the household.

Initial weight for the RDD participants who were from the LowEd area and for the RTS participants= $k/(1 + A) = k/1.55$.

It should be noted that the CLSA conducted RTS but not RDD in NS and conducted RDD but not RTS in PEI. For initial weights construction, we added the NS RTS participants to the NS LowEd HR participants and gave them an initial weight according to their age-sex group. We used k as the initial RDD weight in PEI.

7.2.2. Comprehensive Cohort:

HR mail-outs: We did not assign initial weights to the Comprehensive HR mail-out participants, effectively taking their initial weights to be constant.

Telephone Sampling (TS): The initial weight was assigned as equal to k for the TS participants. We checked the distribution of k by province for both RDD and RTS; we capped k at 3 if $k > 3$.

7.3. Calibration of Initial Weights

7.3.1. Calibration of Tracking Cohort Initial Weights:

Inflation weights for the Tracking cohort were constructed by a calibration of the initial weights to provincial totals for each DCS and Non-DCS area, highest level of education achieved, sex and age group.

The CLSA Tracking (individual) education variable ED_UDR11_TRM categories were grouped into four categories under the name of “individual education”: Grade 11-13 or lower education (QC: Secondary V; NL: 2nd to 4th year of secondary) grouped as “Low Education”; secondary school graduate, no post-secondary education, some post-secondary education, trade certificate or diploma from a vocational school or apprenticeship training and other post-secondary education as “Medium Education”; non-university certificate or diploma from a community college, CEGEP, etc. and university certificate below bachelor’s level as “Higher Education lower”; bachelor’s degree and university degree or certificate above bachelor’s degree as “Higher Education upper”.

The CCHS, TS, and HR mail-out participants and their initial weights were put together, and we created a three-level sample frame variable “sample_frame_3_level”: HR_mailout (HR+targeted_HR), TS (RDD+RTS), and CCHS.

We created another three-level sample frame variable “sample_frame_3_level_2” which reflects also the changes “from NS-RTS to NS-HR_mailout” and “from BC-HR_lowEd to BC-TS”.

There were 83 Tracking participants with missing individual education information. We did hot-deck imputation for the 83 missing values using the “surveyimpute” procedure with the variables “sample_frame_3_level_2”, “province”, “DCS_area (DCS/Non-DCS)”, “sex”, and “age” in the cells statement and “individual education” in var statement. Imputed missing values were flagged; please note that the imputed education variable is only for the purpose of weights construction and imputed values will not be released.

We considered ON to consist of DCS1 (Hamilton), DCS2 (Ottawa) and Non-DCS; QC to consist of DCS1 (McGill), DCS2 (Sherbrooke), and Non-DCS. In BC, the areas for the two DCSs in Vancouver were not separated for purposes of calibration benchmarks; thus, UBC and SFU were considered as DCS1 and Victoria as DCS2.

We selected variables “province”, “age groups (45-48, 49-54, 55-64, 65-74, and 75-85)”, “sex”, individual DCS areas and provincial Non-DCS (AB Calgary, AB Non-DCS, BC SFU_UBC, BC Victoria, BC Non-DCS, MB Manitoba DCS, MB_Non-DCS, NB Non-DCS, NL Memorial, NL Non-DCS, NS Dalhousie, NS Non-DCS, ON Hamilton, ON Ottawa, ON Non-DCS, PE Non-DCS, QC Mc_Gill, QC Sherbrooke, QC Non-DCS, SK Non-DCS), “education (Low, medium, high lower, high upper)” to define our calibration benchmarks.

For the purpose of calibration, we computed the NHS 2011 (weighted) totals within each province crossed with individual DCS (and provincial Non-DCS) crossed with individual education (low, medium, higher lower, higher upper) by sex crossed with age groups (45-48, 49-54, 55-64, 65-74, and 75-85). This is called the NHS table given in Table 18. Note that the age decade 45-54 was split into two intervals, because the sampling rate tended to be appreciably lower for 45-48 than for 49-54.

A raking algorithm was used for the calibration. We used different sets of calibration benchmarks to adjust the basic design weights (initial weights) until the distribution of the weighted sample aligned with the population distribution (NHS 2011) for those variables. Those adjusted weights were accepted as the inflation weights.

First, within each frame sample, and then within each province crossed with sex crossed with age decade, we rescaled the initial weights to sum to sample size, i.e. to have mean 1 and called these the Rescaled InitiaL (RIL) weights.

Secondly, we calibrated the RIL weights with respect to the row sums of the NHS table. That is, in each province we calibrated with respect to DCS-Non-DCS crossed with education. These calibrated weights became inflation weights and were called CIN1 weights.

Thirdly, we checked CIN1 weights to see whether, because of outlier values, there were any cells that should be collapsed.

Fourthly, we calibrated the CIN1 weights with respect to the sex-age groups of the NHS table (Table 18) within each province. Note that the lowest age decade is divided into 45-48 and 49-54 because of the very different sampling fractions in those two categories. We called the new calibrated weights CIN2 weights.

In the fifth step, we checked the CIN2 weights for outlier values and also to see whether they provided an estimated age-sex distribution that was approximately correct for DCS areas. We modified some CIN2 weights trimming large outliers, to produce CIN2_modified weights.

In the sixth step, we calibrated the CIN2_modified weights with respect to sex crossed with the DCS variable within province, to produce CIN3 calibrated weights. We modified some CIN3 weights trimming the smallest outliers, to produce CIN3_altered weights.

We repeated the steps, starting from the second one: the CIN3_altered weights were recalibrated in each province with respect to DCS-Non-DCS crossed with education and called CIN3_altered_calib_to_CIN1 weights. Then, we recalibrated the CIN3_altered_calib_to_CIN1 weights with respect to the sex-age groups within each province, and called the new weights CIN2_II. We trimmed some CIN2_II weights, and reached CIN2_II_trimmed weights. Then CIN2_II_trimmed weights were re-calibrated within province by sex crossed with 5 age groups, and called CIN2_II_trimmed_calib_to_CIN2.

Proportions using the final weights were seen to have approximated the NHS 2011 proportions well on all three dimensions. It was judged that it was not necessary to continue the process.

Here is an example showing the steps of how we calculated the inflation weights Tracking cohort using a real participant in the Tracking cohort, person A:

Rescaled Initial Weights (RIL):

For person A in CCHS-AB-Female-45-54 stratum (stratum a)

np=30 (total number of participants in stratum a)

total_basic_weights=29,794.78 (sum of basic weights for 30 participants in stratum a) basic_weight=775.95 (basic weight for person A)

RIL=$(basic\_weight \times np)/total\_basic\_weights = (775.95 \times 30)/29{,}794.78 = \mathbf{0.781}$

CIN1 weights:

For person A in AB-Calgary DCS-Higher Education upper stratum (stratum b)

sum_of_NHS_weights=110,750 (NHS weight total in stratum b)

sum_of_RIL_prDCSed=153.53 (RIL total for 148 participants in stratum b)

RIL for the person=0.781

CIN1=$(110{,}750 \times 0.781)/153.53 = \mathbf{563.59}$

CIN2 weights:

For person A in AB-female-49-54 stratum (stratum c)

sum_of_NHS_weights_II=164,370 (NHS weight total in stratum c)

sum_of_CIN1_prSEXage=123,992.64 (CIN1 total for 228 participants in stratum c)

CIN1 for the person=563.59

CIN2=$(164{,}370 \times 563.59)/123{,}992.64 = \mathbf{747.12}$

We modified 19 CIN2 weights trimming large outliers: CIN2_modified

CIN3 weights:

For person A in AB-female-Calgary (stratum d)

sum_of_NHS_weights_III=202,875 (NHS weight total in stratum d)

sum_of_CIN2_modified_prSEXdcs=214,454.77 (CIN2_modified total for 195 participants in stratum d)

CIN2_modified for the person=747.12

CIN3=$(202{,}875 \times 747.12)/214{,}454.77 = \mathbf{706.78}$

We modified 58 CIN3 weights trimming the smallest outliers: CIN3_altered

CIN3_altered_calib_to_CIN1 weights:

For person A in AB-Calgary DCS-Higher Education upper stratum (stratum b)

sum_of_NHS_weights=110,750 (NHS weight total in stratum b)

sum_of_CIN3_altered_prDCSed=123,975.51 (CIN3_altered total for 148 participants in stratum b)

CIN3_altered for the person=706.78

CIN3_altered_calib_to_CIN1=$(110{,}750 \times 706.78)/123{,}975.51 = \mathbf{631.38}$

CIN2_II weights:

For person A in AB-female-49-54 stratum (stratum c)

sum_of_NHS_weights_II=164,370 (NHS weight total in stratum c)

sum_of_CIN3_altered_calib_to_CIN1_prSEXage=161,464.36 (CIN3_altered_calib_to_CIN1 total for 228 participants in stratum c)

CIN1_altered_calib_to_CIN1 for the person=631.38

CIN2_II=$(164{,}370 \times 631.38)/161{,}464.36 = \mathbf{642.74}$

We trimmed 12 CIN2_II weights: CIN2_II_trimmed.

Then CIN2_II_trimmed weights were re-calibrated within province by sex crossed with 5 age groups:

CIN2_II_trimmed_calib_to_CIN2.

It was judged that it was not necessary to continue the process.

CIN2_II_trimmed_calib_to_CIN2 weights were accepted as inflation weights for the Tracking cohort and they add up to 13,655,060.

7.3.2. Calibration of Comprehensive Cohort Initial Weights:

Comprehensive TS and HR mail-out participants and their initial weights were put together.

The CLSA Comprehensive (individual) education variable *ED_UDR11_COM* categories were grouped into “Low Education”, “Medium Education”, “Higher Education, upper”, and “Higher Education, lower” under the name of “individual education” as we did for the CLSA Tracking education variable *ED_UDR11_TRM*.

We created a two-level sample frame “sample_frame_2_level”: a. HR_mailout (HR+HR_lowEd), b. TS (RDD+RTS).

We considered ON consisting of DCS1 (Hamilton), DCS2 (Ottawa) and Non-DCS; QC consisting of DCS1 (McGill), DCS2 (Sherbrooke), and Non-DCS. In BC, the areas for the two DCSs in Vancouver, were not separated for purposes of calibration benchmarks; so, UBC and SFU were considered as DCS1 and Victoria as DCS2.

There were 50 Comprehensive participants with missing individual education information. We did hot-deck imputation for the 50 missing values using the “surveyimpute” procedure with the variables “sample_frame_2_level”, “province”, “individual DCS”, “sex”, and “age” in the cells statement and “individual education” in var statement. Imputed missing values were flagged; please note that the imputed education variable is only for the purpose of weights construction and imputed values will not be released. In the Tracking cohort, we used DCS/Non-DCS information while imputing the missing values; here we used the individual DCS.

For the purpose of calibration, we used the NHS 2011 dataset in DCS areas. After excluding Non-DCS areas from the NHS 2011 dataset, we had a total WEIGHT of 3,812,085.

We computed the NHS (weighted) proportions within individual DCS crossed with individual education (low, medium, higher upper, higher lower from the HCDD_7V variable in the NHS) by sex crossed with age groups (45-48, 49-54, 55-64, 65-74, and 75-85).

The same raking approach that was used with the Tracking sample was applied to the Comprehensive sample. We used calibration benchmarks to adjust the design weights until the distribution of the weights aligned with population distribution (NHS 2011) in DCS areas for those variables.

For the HR mail-out participants, we directly assigned 1 as their RIL weights; for the TS participants, the RIL weights were obtained by rescaling the initial weights to sum to sample size within DCS crossed with sex crossed with age decade.

Here is an example showing the steps of how we calculated the inflation weights and analytic weights for the Comprehensive cohort using a real participant in the Comprehensive cohort, person B:

Rescaled Initial Weights (RIL):

For person B in TS-Calgary-Famale-45-54 stratum (stratum a)

n_of_CLSA_part_in_DCSagesex=384 (total number of participants in stratum a)

total_basic_weights=612 (sum of basic weights for 384 participants in stratum a)

basic_weight=2 (basic weight for person B)

RIL= $(basic\_weight \times n\_of\_CLSA\_part\_in\_DCSagesex)/total\_basic\_weights = (2 \times 384)/612 = \mathbf{1.2549}$

CIN1 weights:

For person B in AB-Calgary DCS-Higher Education upper stratum (stratum b)

sum_of_NHS_weights=110,750 (NHS weight total in stratum b)

sum_of_RIL_prDCSed=1,506.644 (RIL total for 1504 participants in stratum b)

RIL for the person=1.2549

CIN1=$(110{,}750 \times 1.2549)/1{,}506.644 = \mathbf{92.2449}$

CIN2 weights:

For person B in Calgary-female-49-54 stratum (stratum c)

sum_of_NHS_weights_II=53,710 (NHS weight total in stratum c)

sum_of_CIN1_dcsSEXage=35,399.866 (CIN1 total for 279 participants in stratum c)

CIN1 for the person=92.2449

CIN2=$(53{,}710 \times 92.2449)/35{,}399.866 = \mathbf{139.9574}$

We modified 232 CIN2 weights trimming the outliers: CIN2_modified

CIN2_modified_calib_to_CIN1 weights:

For person B in AB-Calgary DCS-Higher Education upper stratum (stratum b)

sum_of_NHS_weights=110,750 (NHS weight total in stratum b)

sum_of_CIN2_modified_prDCSed=125,483.150 (CIN2_modified total for 1504 participants in stratum b)

CIN2_modified for the person=139.9575

CIN2_modified_calib_to_CIN1=$(110{,}750 \times 139.9575)/125{,}483.150 = \mathbf{123.5249}$

CIN2_II weights:

For person B in Calgary-female-49-54 stratum (stratum c)

sum_of_NHS_weights_II=53,710 (NHS weight total in stratum c)

sum_of_ CIN2_modified_calib_to_CIN1_dcsSEXage=52,188.1713 (CIN2_modified_calib_to_CIN1 total for 279 participants in stratum c)

CIN2_modified_calib_to_CIN1 for the person=123.5249

CIN2_II=$(53{,}710 \times 123.5249)/52{,}188.1713 = \mathbf{127.13}$

CIN2_II is the inflation weight variable for the Comprehensive cohort and they add up to 3,812,085.

7.4. Calculation of Analytic Weights

Tracking inflation weights were further rescaled linearly to add to the number of participants within each province to provide analytic weights, and Comprehensive inflation weights were rescaled linearly to add to the number of participants within individual DCS. The analytic weights add up to 21,241 for the Tracking cohort and 30,097 for the Comprehensive cohort.

7.4.1. Rationale for Rescaling within Provinces

The inflation weights are to be used for *descriptive purposes*, such as estimating population averages of continuous variables, or numbers and proportions of population members with characteristics of interest. They are computed so that weighted sample sums will estimate population totals without bias. For *analytical purposes*, however, some different considerations apply. Regression and logistic regression analyses may be designed to estimate the relationships among variables, not so much for the population at hand as for a hypothetical population of like people, represented by the sample. The estimates of the coefficients or other parameters are found by solving estimating equations which are combinations of terms based on model residuals. For example, if we want to study self-reported BMI as a predictor of rheumatoid arthritis, using data from an entire moderate sized population, the *census estimating equation* for the logistic regression would look like this:

$$\sum_{i=1}^{N} X_i \left(y_i - g\left(X_i^T \beta\right)\right) = 0, \qquad (1)$$

where y_i is 1 if i has rheumatoid arthritis and 0 otherwise; X_i is a column vector of k explanatory variables or covariates including a suitable transformation of BMI; *T* denotes transpose; β is a column vector of k regression coefficients; and $g(z) = \frac{e^z}{1+e^z}$. Under the logistic regression model, each term of (1) has expectation 0, and solving (1) yields β_N, the census estimate of β. If the model is true, then because *N* is very large, β_N will be very close to β. If the model is not true, but the correct model is about the same for all provinces, β_N may still be a useful quantity, summarizing the relationship of *X* and *y* in the population.

The *sample-based estimating equation* that is approximately sampling-design-unbiased for the census estimating equation (1) is this:

$$\sum_{i \in s} w_i X_i \left(y_i - g(X_i^T \beta)\right) = 0, \qquad (2)$$

where s is the set of individuals in the sample.

The solution of (2) is a sampling-design-consistent estimator of β_N. Also, if in the sense of the model, the weight w_i given X_i is independent of or uncorrelated with the residual $y_i - g(X_i^T \beta)$, then the solution of (2) is a consistent estimator of β in the sense of the model. But the same would be true for any equation like (2) with a different choice of weights, as long as the expectation of each term is still 0. If the assumed logistic regression model is true, and the weights w_i are highly variable over the population, solving (2) to estimate β will be very inefficient compared to the same equation with a choice of weights which is closer to uniform.

In the CLSA design, the sampling rates are smaller in the larger provinces, and within a province, the sampling rates in the DCS areas are smaller (for the pooled sample) than in the Non-DCS areas. Thus, the use of the inflation weights in (2) will mean that if the model is correct, the estimation of β will be inefficient. On the other hand, if the model is only an approximation, in the sense that β varies from province to province, then the terms in the larger provinces being dominant, the estimated value of β would be the value appropriate to the larger provinces. In either case, it may be more satisfactory to use less variable weights for this analytical purpose, and at the same time include province (or province crossed with DCS/Non-DCS) in the model, as well as potentially a term for interaction of province or area and BMI.

7.4.2. Construction

The so-called analytic weights supplied with the Tracking cohort data [WGHTS_ANALYTIC_TRM] are proportional to the inflation weights but rescaled to sum to the sample size within each province, so that their mean value is 1 within each province.

Once this rescaling is done, producing analytic weights $\widetilde{w}_i$, then the sample-based estimating equation

$$\sum_{i \in s} \widetilde{w}_i X_i \left(y_i - g\left(X_i^T \beta\right)\right) = 0 \qquad (3)$$

is approximately unbiased for the new census estimating equation

$$\sum_h \frac{1}{N_h} \sum_{i=1}^{N_h} X_i \left(y_i - g\left(X_i^T \beta\right) \right) = 0 \qquad (4)$$

where N_h is the number in the study population in province h. The solution of (4) would be a new census parameter $\tilde{\beta}_N$, close to an average of the values summarizing the relationships in the provinces. The estimating equation (3) will provide a design-consistent estimator of this new census parameter, which is a model-consistent estimator of β if the model is correct; the estimating equation (3) gives a more efficient estimator than would be provided by estimating equation (2). Here again, the estimating equation (3) is model-unbiased if the weight given X_i is independent of the residual.

The analytic weights supplied with the Comprehensive cohort data [WGHTS_ANALYTIC_COM] are proportional to the inflation weights but rescaled to sum to sample size within the individual DCS part of each province, so that their mean value is 1 within that area.

The analytic weights supplied with the pooled data [WGHTS_ANALYTIC_CLSAM] are proportional to the inflation weights for the pooled data but rescaled to sum to sample size within the individual DCS and Non-DCS part of each province.

7.4.3. Variables to Include in the Model

Users should add the weighting variables sex, age group (45-54, 55-64, 65-74, and 75-85), individual education (low, medium, high lower, high upper) to their analyses very routinely, because those are obvious potential confounders of any association, and also consider adding province for Tracking and province crossed with individual DCS areas for Comprehensive and pooled cohort. Not only are these geographic categories named as strata, but they were also used in calibration of the weights. By including them as covariates we try to ensure that the residuals in the regression or their analogues for other generalized linear models are at least approximately independent of the geographic part of the sampling design.

7.4.4. Why Using Weights is Recommended in Modeling

The main reason for the standard advice to use the weights in analysis has to do with possible informativeness of the sampling design and/or the response-nonresponse process. That is, it could be the case that the distribution of the residual $y_i - g\left(X_i^T \beta\right)$, conditional on X_i, depends on whether the unit i is sampled, to the point that under the combination of model and design, the term $X_i(y_i - g(X_i^T \beta))$ does not have expectation 0, given that i is in the final sample. For example, suppose marital status has an influence on the relationship between BMI and rheumatoid arthritis, but is not included in the model. Marital status influences the individual's inclusion probability, which for the RDD part of the sample is lower, the more eligible individuals there are in the household. In that case, inclusion in the sample is informative about the relationship of interest. For another example, the design has over-sampled people in low education areas, which might also be areas where environmental factors could contribute to chronic conditions. Here also, being in the sample might alter the statistical relationship between BMI and rheumatoid arthritis.

Under these kinds of circumstances, the terms in (3) will also not have model expectation 0, but because of the presence of the weights, (3) will still be an unbiased estimator of the province population level estimating equation (4), and hence the solution of (3) will be a design-consistent estimator of $\tilde{\beta}_N$.

Software is available to handle most kinds of analyses where the methodology for using weights has been developed and standardized. Packages with survey data features include R, SAS, SPSS, Stata, SUDAAN and Mplus.

7.5. When Not to Use the Weights

For a description of the sample itself, rather than the study population or a hypothetical population behind it, the weights are not used.

The weights might also not be used for an analysis where the object of inference is not readily expressible in terms of estimating equations of the same kind of form as (1) or (4).

An example might be the use of survey data in some kinds of spatial analysis. In such cases it is best to consult with statisticians.

7.6. Sample Weights for the Pooled Data

The purpose of pooling the Tracking and Comprehensive sample would be to permit analyses of variables that are common to the two questionnaires. It would be desirable that results of analyses with Tracking cohort alone and with the pooled sample be comparable as far as possible, but users should be aware that they may not be the same.

For the pooled sample, we combined the Tracking cohort analytic weights and Comprehensive cohort analytic weights. Because the Tracking cohort analytic weights were rescaled within province, the mean value of the pooled analytic weights within individual DCS was not 1. We decided to create a new (second) analytic weight variable to be used in the pooled cohort, rescaling the Tracking inflation weights to sum to sample size within each individual DCS and provincial Non-DCS area. This new analytic weight variable is not going to be used/released with the Tracking cohort; it is only for the pooled cohort.

The union of this new Tracking analytic weight and the Comprehensive analytic weight sums to sample size within each DCS and provincial Non-DCS area, and the mean value of the pooled analytic weights within DCS is 1.

The pooled analytic weights add up to 51,338, the total number of CLSA participants.

We then calculated the pooled inflation weights recalibrating the pooled analytic weights to the NHS total by individual DCS (and provincial Non-DCS) crossed with sex crossed with age group (5 groups). The pooled inflation weights add up to 13,655,060.

The pooled analytic weights of Comprehensive cohort participants are the same as their Comprehensive analytic weights, but the pooled analytic weights of Tracking cohort participants are different from their original Tracking weights. Please note that because there is no DCS in NB, PEI, SK, the Tracking participants' pooled analytic weights are the same as their Tracking analytic weights; so are their Tracking inflation weights and pooled inflation weights.

8. CALCULATION of FOLLOW-UP 1 SAMPLE (INFLATION) WEIGHTS

The first follow-up (FUP1) was completed (2015-2018) with 95.6% retention (4.4% withdrawal since baseline); 17,326 Tracking participants and 27,765 Comprehensive participants.

8.1. Source of Benchmarks for Calibration of Follow-up 1 Cross-sectional Weights: Census 2016

For the follow-up 1 cross-sectional weights, we followed the same method used in the calculation of new (updated) baseline sample weights (see section 7, page 10), but unlike the baseline we used Census 2016 as our population reference for calibration purposes.

The 2016 Census microdata file contains 8,651,677 observations and 663 variables. The target population consists of all persons who completed form 2A-L or 2A-R in the 2016 Census. The file does not include people living in institutions, Canadian citizens temporarily living in other countries, full-time members of Canadian Forces stationed outside Canada, persons living in institutional collective dwellings such as hospitals, nursing homes and penitentiaries, persons living in non-institutional collective dwellings such as work camps, hotels and motels, and student residences. Each record in the Census microdata file represents a person level response and is the primary level on which the Census microdata file is based.

After excluding households on First Nations reserves and Aboriginal Inuit areas of residence, people who were not permanent residents, people who lived in the 3 territories, and people younger than 49 or older than 89 from the Census 2016 dataset, we ended up with 3,178,308 people with a total weight of 13,130,030.00, summing the Census 2016 inflation weights (Comp_W2).

8.2. Calibration of Initial Weights

8.2.1. Calibration of Tracking Cohort Initial Weights

Predicted probability of death by FUP1 (pred_death) and predicted probability of retention at FUP1 (pred_retention) for each Tracking participant who completed FUP1 were calculated from logistic regression models using CLSA baseline variables sex, age group, individual education and perceived health as predictor variables. Retention adjusted weights were calculated multiplying baseline RIL values with (1-pred_death)/pred_retention and considered as initial FUP1 weights.

FUP1 inflation weights for the Tracking cohort were constructed by a calibration of the retention adjusted weights to provincial totals for each DCS and Non-DCS area, highest level of education achieved, sex and age group.

For the purpose of calibration, we computed the Census 2016 (weighted) totals within each province crossed with individual DCS (and provincial Non-DCS) crossed with individual education (low, medium, higher lower, higher upper) by sex crossed with age groups (49-54, 55-64, 65-74, and 75-84, 85-89).

The same raking approach that was used with the baseline Tracking cohort was applied to the FUP1 Tracking cohort for the calibration. We used different sets of calibration benchmarks to adjust the initial FUP1 weights (retention adjusted weights) until the distribution of the weighted sample aligned with the population distribution (Census 2016) for those variables. Those adjusted weights were accepted as the inflation weights.

Inflation weights were assigned to zero for the Tracking participants who moved to a territory. For the participants who moved to another province (different from the baseline province), inflation weights were calculated using the new province.

8.2.2. Calibration of Comprehensive Cohort Initial Weights

Predicted probability of death by FUP1 (pred_death) and predicted probability of retention at FUP1 (pred_retention) for each Comprehensive participant who completed FUP1 were calculated from logistic regression models using baseline variables sex, age group, individual education and perceived health as predictor variables. Retention adjusted weights were calculated multiplying baseline RIL values with (1-pred_death)/pred_retention and considered as initial FUP1 weights.

FUP1 inflation weights for the Comprehensive cohort were constructed by a calibration of the retention adjusted weights to individual DCS totals, highest level of education achieved, sex and age group.

For the purpose of calibration, we used the Census 2016 dataset in DCS areas. After excluding Non-DCS areas from the Census 2016 dataset, we had a total WEIGHT of 3,743,900. We computed the Census 2016 (weighted) proportions within individual DCS crossed with individual education (low, medium, higher lower, higher upper from the HCDD_7V variable in the Census 2016) by sex crossed with age groups (49-54, 55-64, 65-74, and 75-84, 85-89).

The same raking approach that was used with the baseline Comprehensive cohort was applied to FUP1 Comprehensive cohort for the calibration. We used different sets of calibration benchmarks to adjust the initial FUP1 weights (retention adjusted weights) until the distribution of the weighted sample aligned with the population distribution (Census 2016) for those variables. Those adjusted weights were accepted as the inflation weights.

For the Comprehensive participants who moved to a territory or Non-DCS area, inflation weights were calculated using the DCS the participants completed the FUP1 visit. For the participants who moved to another DCS (different from the baseline DCS), inflation weights were calculated using the new DCS.

8.3. Calculation of Follow-up 1 Analytic Weights

FUP1 Tracking inflation weights were further rescaled linearly to add to the number of participants within each province to provide analytic weights, and FUP1 Comprehensive inflation weights were rescaled linearly to add to the number of participants within individual DCS. The FUP1 analytic weights add up to 17,324 (2 participants from Yukon and Nunavut were assigned 0 as FUP1 inflation and analytic weights) for the Tracking cohort and 27,765 for the Comprehensive cohort.

8.4. Follow-up 1 Sample Weights for the Pooled Data

The same approach used to get baseline pooled analytic and inflation weights was followed.

The FUP1 pooled analytic weights add up to 45,089, the total number of CLSA FUP1 participants. The FUP1 pooled inflation weights were calculated recalibrating the pooled analytic weights to the Census 2016 total by individual DCS (and provincial Non-DCS) crossed with sex crossed with age group (5 groups). The pooled inflation weights add up to 13,130,130.

9. CALCULATION of FOLLOW-UP 2 SAMPLE (INFLATION) WEIGHTS

The second follow-up (FUP2) was completed (2018-2021) with 92.8% retention (7.2% withdrawal since baseline); 15,211 Tracking participants and 25,605 Comprehensive participants.

9.1. Source of Benchmarks for Calibration of Follow-up 2 Cross-sectional Weights: Census 2016

For the follow-up 2 cross-sectional weights, we followed the same method used in the calculation of new (updated) baseline sample weights (see section 7, page 10) and follow-up 1 sample weights (see section 8, page 24).

We used Census 2016 as our population reference for calibration purposes as we did for the calibration of follow-up 1 sample weights, but this time we excluded people younger than 55 and older than 89 from the Census 2016 dataset; we ended up with 2,435,489 people with a total weight of 10,059,354.18, summing the Census 2016 inflation weights (Comp_W2).

9.2. Calibration of Initial Weights

9.2.1. Calibration of Tracking Cohort Initial Weights

Predicted probability of death by FUP2 (pred_death) and predicted probability of retention at FUP2 (pred_retention) for each Tracking participant who completed FUP2 were calculated from logistic regression models using CLSA baseline variables sex, age group, individual education and perceived health as predictor variables. Retention adjusted weights were calculated multiplying baseline RIL values with (1-pred_death)/pred_retention and considered as initial FUP2 weights.

FUP2 inflation weights for the Tracking cohort were constructed by a calibration of the retention adjusted weights to provincial totals for each DCS and Non-DCS area, highest level of education achieved, sex and age group.

For the purpose of calibration, we computed the Census 2016 (weighted) totals within each province crossed with individual DCS (and provincial Non-DCS) crossed with individual education (low, medium, higher lower, higher upper) by sex crossed with age groups (55-64, 65-74, and 75-84, 85-89).

The same raking approach that was used with the baseline/FUP1 Tracking cohort was applied to the FUP2 Tracking cohort for the calibration. We used different sets of calibration benchmarks to adjust the initial FUP2 weights (retention adjusted weights) until the distribution of the weighted sample aligned with the population distribution (Census 2016) for those variables. Those adjusted weights were accepted as the inflation weights.

Inflation weights were assigned to zero for the Tracking participants who moved to a territory. For the participants who moved to another province (different from the baseline province), inflation weights were calculated using the new province. For the proxy participants who lived in institutions during FUP2 interviews, inflation weights were assigned to zero.

9.2.2. Calibration of Comprehensive Cohort Initial Weights

Predicted probability of death by FUP2 (pred_death) and predicted probability of retention at FUP2 (pred_retention) for each Comprehensive participant who completed FUP2 were calculated from logistic regression models using baseline variables sex, age group, individual education and perceived health as predictor variables. Retention adjusted weights were calculated multiplying baseline RIL values with (1-pred_death)/pred_retention and considered as initial FUP2 weights.

FUP2 inflation weights for the Comprehensive cohort were constructed by a calibration of the retention adjusted weights to individual DCS totals, highest level of education achieved, sex and age group.

For the purpose of calibration, we used the Census 2016 dataset in DCS areas. After excluding Non-DCS areas from the Census 2016 dataset, we had a total WEIGHT of 2,833,140. We computed the Census 2016 (weighted) proportions within individual DCS crossed with individual education (low, medium, higher lower, higher upper from the HCDD_7V variable in the Census 2016) by sex crossed with age groups (55-64, 65-74, and 75-84, 85-89).

The same raking approach that was used with the baseline/FUP1 Comprehensive cohort was applied to FUP2 Comprehensive cohort for the calibration. We used different sets of calibration benchmarks to adjust the initial FUP2 weights (retention adjusted weights) until the distribution of the weighted sample aligned with the population distribution (Census 2016) for those variables. Those adjusted weights were accepted as the inflation weights.

For the Comprehensive participants who moved to a territory or Non-DCS area, inflation weights were calculated using the DCS the participants completed the FUP2 visit. For the participants who moved to another DCS (different from the baseline DCS), inflation weights were calculated using the new DCS. For the proxy participants who lived in institutions during FUP2 interviews, inflation weights were assigned to zero.

9.3. Calculation of Follow-up 2 Analytic Weights

FUP2 Tracking inflation weights were further rescaled linearly to add to the number of participants within each province to provide analytic weights, and FUP2 Comprehensive inflation weights were rescaled linearly to add to the number of participants within individual DCS. The FUP2 analytic weights add up to 15,204 (1 participant from Nunavut and 6 proxy participants were assigned 0 as FUP2 inflation and analytic weights) for the Tracking cohort and 25,596 for the Comprehensive cohort (9 proxy participants were assigned 0 as FUP2 inflation and analytic weights).

9.4. Follow-up 2 Sample Weights for the Pooled Data

The same approach used to get baseline/FUP1 pooled analytic and inflation weights was followed.
The FUP2 pooled analytic weights add up to 40,800. The FUP2 pooled inflation weights were calculated recalibrating the pooled analytic weights to the Census 2016 total by individual DCS (and provincial Non-DCS) crossed with sex crossed with four-level age group (55-64, 65-74, 75-84, 85-89). The pooled inflation weights add up to 10,059,510.

10. CALCULATION of "BASELINE to FUP1" WEIGHTS

10.1. Calculation of Tracking Longitudinal Inflation Weights

The predicted probability of not being lost to follow-up by FUP1 (includes participants who provided FUP1 data and those who died between baseline and FUP1) was calculated for each Tracking participant who completed baseline from a fitted logistic regression model using CLSA baseline variables sex, age group, individual education and perceived health as predictor variables. Baseline inflation weights for the participants who were not lost to follow-up by FUP1 were divided by the predicted probabilities of not being lost to follow-up, and the results were considered as initial longitudinal weights.

Initial longitudinal weights were calibrated to the 2011 National Household Survey (weighted) totals for each provincial DCS and Non-DCS, individual education, sex and age group until the distribution of the weighted sample aligned with the population distribution for those variables. Those adjusted weights were accepted as the Tracking Baseline to FUP1 longitudinal inflation weights.

Two sets of analytic weights were calculated: (a) The first set was calculated for participants not being lost to follow-up and can be used to model as an outcome by FUP1. They are proportional to the longitudinal inflation weights but rescaled to sum to the sample size within each province, so that their mean value is 1 within each province. (b) The second set was calculated for participants who have not died and have provided data at FUP1 and can be used to study health of individuals conditional on their not having died by the time of FUP1. They are proportional to longitudinal inflation weights but rescaled to sum to the sample size within each province, so that their mean value is 1 within each province.

10.2. Calculation of Comprehensive Longitudinal Inflation Weights

The predicted probability of not being lost to follow-up by FUP1 (includes participants who provided FUP1 data and those who died between baseline and FUP1) was calculated for each Comprehensive participant who completed baseline from a logistic regression model using CLSA baseline variables sex, age group, individual education and perceived health as predictor variables. Baseline inflation weights for the participants who were not lost to follow-up by FUP1 were divided by the predicted probabilities of not being lost to follow-up, and the results were considered as initial longitudinal weights.

Initial longitudinal weights were calibrated to 2011 National Household Survey (weighted) totals for each DCS, individual education, sex and age group until the distribution of the weighted sample aligned with the population distribution for those variables. Those adjusted weights were accepted as the Comprehensive Baseline to FUP1 longitudinal inflation weights.

Two sets of analytic weights were calculated: (a) The first set was calculated for participants not being lost to follow-up and can be used to model death by the time of FUP1. They are proportional to the longitudinal inflation weights but rescaled to sum to the sample size within each DCS, so that their mean value is 1 within each DCS. (b) The second set was calculated for participants who have not died and have provided data at FUP1 and can be used to study health of individuals conditional on their not having died by the time of FUP1. They are proportional to longitudinal inflation weights but rescaled to sum to the sample size within each DCS, so that their mean value is 1 within each DCS.

10.3. Calculation of Pooled Longitudinal Inflation Weights

For the pooled sample, we combined longitudinal analytic weights from the Tracking and Comprehensive cohorts. Because the Tracking cohort analytic weights were rescaled within province, the mean value of the pooled analytic weights within individual DCSs was not 1. We decided to create two new analytic weight variables to be used in the pooled cohort, rescaling the Tracking longitudinal inflation weights to sum to sample size within each individual DCS and provincial Non-DCS area: (a) for participants not lost to follow-up by the time of FUP1 and (b) for participants who have not died and have provided data at FUP1. These new analytic weight variables are not going to be released for the Tracking cohort; they are only for the pooled cohort.

The union of these new Tracking analytic weights and the Comprehensive analytic weights sums to the sample size within each DCS and provincial Non-DCS area, and the mean value of the pooled analytic weights within DCS is 1.

We then calculated the pooled longitudinal inflation weights recalibrating the pooled analytic weights for participants not t to follow-up by the time of FUP1 to the 2011 National Household Survey total by individual DCS (and provincial Non-DCS) crossed with sex crossed with age group (5 groups). The pooled longitudinal inflation weights add up to 13,655,060.

The pooled analytic weights for Comprehensive cohort participants are the same as their Comprehensive analytic weights, but the pooled analytic weights for Tracking cohort participants are different from their original Tracking weights. Please note that because there is no DCS in NB, PEI, or SK, the Tracking participants' pooled analytic weights are the same as their Tracking analytic weights; so are their Tracking longitudinal inflation weights and pooled longitudinal inflation weights.

TABLES FOR BASELINE

Table 1. Actual and Target Number of the CLSA Tracking Participants

Province*	Age Group	Sex	Target Number of Participants in Stratum	Actual Number of Participants in Stratum
AB	45-54	Female	306	339
AB	45-54	Male	306	311
AB	55-64	Female	306	348
AB	55-64	Male	306	314
AB	65-74	Female	189	204
AB	65-74	Male	189	205
AB	75-85	Female	189	190
AB	75-85	Male	189	196
BC	45-54	Female	379	407
BC	45-54	Male	379	360
BC	55-64	Female	379	431
BC	55-64	Male	379	403
BC	65-74	Female	234	271
BC	65-74	Male	234	255
BC	75-85	Female	234	255
BC	75-85	Male	234	238
MB	45-54	Female	212	228
MB	45-54	Male	212	224
MB	55-64	Female	212	240
MB	55-64	Male	212	216
MB	65-74	Female	141	141
MB	65-74	Male	141	149
MB	75-85	Female	141	151
MB	75-85	Male	141	135
NB	45-54	Female	190	210
NB	45-54	Male	190	195
NB	55-64	Female	190	212
NB	55-64	Male	190	201
NB	65-74	Female	127	138
NB	65-74	Male	127	143
NB	75-85	Female	127	131
NB	75-85	Male	127	129
NL	45-54	Female	173	190
NL	45-54	Male	173	173
NL	55-64	Female	173	189
NL	55-64	Male	173	196
NL	65-74	Female	125	126
NL	65-74	Male	125	128
NL	75-85	Female	125	123
NL	75-85	Male	125	128

Province	Age Group	Sex	Target Number of Participants in Stratum	Actual Number of Participants in Stratum
NS	45-54	Female	205	227
NS	45-54	Male	205	223
NS	55-64	Female	205	251
NS	55-64	Male	205	233
NS	65-74	Female	137	167
NS	65-74	Male	137	170
NS	75-85	Female	137	131
NS	75-85	Male	137	151
ON	45-54	Female	658	694
ON	45-54	Male	658	674
ON	55-64	Female	658	755
ON	55-64	Male	658	722
ON	65-74	Female	439	518
ON	65-74	Male	439	460
ON	75-85	Female	439	459
ON	75-85	Male	439	440
PE	45-54	Female	150	165
PE	45-54	Male	150	160
PE	55-64	Female	150	165
PE	55-64	Male	150	151
PE	65-74	Female	125	127
PE	65-74	Male	125	127
PE	75-85	Female	125	121
PE	75-85	Male	125	127
QC	45-54	Female	525	581
QC	45-54	Male	525	526
QC	55-64	Female	525	577
QC	55-64	Male	525	575
QC	65-74	Female	350	349
QC	65-74	Male	350	366
QC	75-85	Female	350	314
QC	75-85	Male	350	320
SK	45-54	Female	202	217
SK	45-54	Male	202	189
SK	55-64	Female	202	221
SK	55-64	Male	202	215
SK	65-74	Female	134	146
SK	65-74	Male	134	144
SK	75-85	Female	134	129
SK	75-85	Male	134	131

*** AB=Alberta, BC=British Columbia, MB=Manitoba, NB=New Brunswick, NL=Newfoundland and Labrador, NS=Nova Scotia, ON=Ontario, PE=Prince Edward Island, QC=Quebec, SK=Saskatchewan.**

Table 2. Actual and Target Number of the CLSA Comprehensive Participants

Province	Age Group	Sex	Target Number of Participants in Stratum	Actual Number of Participants in Stratum
AB	45-54	Female	450	384
AB	45-54	Male	450	329
AB	55-64	Female	450	509
AB	55-64	Male	450	492
AB	65-74	Female	300	371
AB	65-74	Male	300	375
AB	75-85	Female	300	253
AB	75-85	Male	300	244
BC	45-54	Female	900	831
BC	45-54	Male	900	782
BC	55-64	Female	900	1030
BC	55-64	Male	900	980
BC	65-74	Female	600	724
BC	65-74	Male	600	737
BC	75-85	Female	600	573
BC	75-85	Male	600	597
MB	45-54	Female	450	415
MB	45-54	Male	450	366
MB	55-64	Female	450	527
MB	55-64	Male	450	511
MB	65-74	Female	300	373
MB	65-74	Male	300	367
MB	75-85	Female	300	279
MB	75-85	Male	300	275
NL	45-54	Female	450	309
NL	45-54	Male	450	274
NL	55-64	Female	450	379
NL	55-64	Male	450	343
NL	65-74	Female	300	265
NL	65-74	Male	300	264
NL	75-85	Female	300	179
NL	75-85	Male	300	201
NS	45-54	Female	450	391
NS	45-54	Male	450	378
NS	55-64	Female	450	499
NS	55-64	Male	450	460
NS	65-74	Female	300	389
NS	65-74	Male	300	424
NS	75-85	Female	300	270
NS	75-85	Male	300	267
ON	45-54	Female	900	803
ON	45-54	Male	900	781
ON	55-64	Female	900	1070
ON	55-64	Male	900	1051
ON	65-74	Female	600	780
ON	65-74	Male	600	788
ON	75-85	Female	600	554
ON	75-85	Male	600	591
QC	45-54	Female	900	792
QC	45-54	Male	900	760
QC	55-64	Female	900	1075
QC	55-64	Male	900	930
QC	65-74	Female	600	786
QC	65-74	Male	600	719
QC	75-85	Female	600	510
QC	75-85	Male	600	491

Table 3. Disposition of Participants from Canadian Community Health Survey (CCHS)

<table>
<tr><td rowspan="4">People aged between 45-85</td><td rowspan="2">Eligible participants who allowed Statistics Canada to pass on their CCHS data to the CLSA</td><td>Survey + Contact</td><td>11,742</td><td rowspan="4">26,248</td></tr>
<tr><td>Survey Only</td><td>8,345</td></tr>
<tr><td colspan="2">Contact Only</td><td>527</td></tr>
<tr><td colspan="2">Neither</td><td>5,634</td></tr>
<tr><td colspan="4">People aged 85+</td><td>4,617</td></tr>
<tr><td colspan="4">TOTAL</td><td>30,865</td></tr>
</table>

Table 4. Exclusion Criteria Implemented by the Provinces for the Mail-outs from Health Registries

Province*	People living in Long-term Care Facilities and Institutions	People living in First Nations Settlements	Living on Crown Lands	People Who are not Canadian Citizen/ Landed Immigrant/Permanent Resident
BC	Excluded	Excluded**	Excluded	
MB	Excluded	Excluded	Excluded	
NB	Excluded	Excluded**		
NL		Excluded	Excluded	
NS	Excluded	Excluded**		
ON	Excluded	Excluded**		
PE	Excluded	Excluded**		
SK	Excluded	Excluded**		Excluded

*** In the Tracking cohort: SK and ON were included only in the first mail-out and BC was included only in the second mail-out; in the Comprehensive cohort: the first mail-out was conducted in MB, NL, NS, ON, and BC, and the second mail-out was conducted in only NS.**

**** For the second mail-out, the CLSA team provided the provinces with the postal codes excluding First Nations settlements. Only MB and NL excluded the First Nations settlements themselves.**

Table 5. CLSA Tracking HR1-Initial Mail-outs by Provincial Health Registries; Number of Letters Sent by Age-Sex Groups

TRACKING COHORT											
Province	**Number of Mail-outs Conducted**	**Year**	**Number of People Mailed Out by Age-Sex Groups**								**TOTAL**
			M 45-54	**F 45-54**	**M 55-64**	**F 55-64**	**M 65-74**	**F 65-74**	**M 75-85**	**F 75-85**	
MB	3	2012	939	835	720	720	410	405	733	700	5,462
SK	1	2012	920	915	780	525	430	405	660	653	5,288
ON	2	2012, 2014	2,354	2,696	2,042	2,743	1,818	2,011	2,200	2,566	18,430
NS	2	2012	1,883	1,173	1,075	763	473	440	845	774	7,426
NB	2	2011, 2012	832	784	666	639	391	357	616	547	4,832
PE	2	2011, 2012	660	645	570	450	470	450	700	687	4,632
NL	2	2012	795	765	635	575	480	435	747	693	5,125
TOTAL											**51,195**

Table 6. CLSA Comprehensive HR1-Initial Mail-outs by Provincial Health Registries; Number of Letters Sent by Age-Sex Groups

COMPREHENSIVE COHORT											
Province	**Number of Mail-outs Conducted**	**Year**	**Number of People Mailed Out by Age-Sex Groups**								**TOTAL**
			M 45-54	**F 45-54**	**M 55-64**	**F 55-64**	**M 65-74**	**F 65-74**	**M 75-85**	**F 75-85**	
MB	3	2012	750	750	750	750	500	500	667	667	5,334
NL	2	2012	450	450	449	450	299	300	2448	2096	6,942
NS	5	2012, 2013, 2014	2,606	1,504	1,520	1,125	750	750	1,081	1,333	10,669
ON*	6	2012	1,500	1,500	1,500	1500	1000	1000	1333	1333	10,666
BC	2	2011, 2012	8,250	0	0	0	0	0	0	0	8,250
TOTAL											**41,861**

Table 7. CLSA Tracking HR2-Mail-outs by Health Registries Targeting Low-Education Areas: Number of Letters Sent by Age-Sex Groups

TRACKING COHORT											
Province	Number of Mail-outs Conducted	Year	Number of People Mailed Out by Age-Sex Groups								TOTAL
			M 45-54	F 45-54	M 55-64	F 55-64	M 65-74	F 65-74	M 75-85	F 75-85	
MB	1	2014	201	99	331	190	69	276	247	251	1,664
BC	1	2014	697	809	637	379	244	325	470	689	4,250
NS	1	2014	1,026	600	517	433	618	464	685	460	4,803
NB	1	2014	834	448	168	134	98	95	331	757	2,865
PE	1	2014	683	385	311	124	376	173	393	497	2,942
NL	1	2014	728	170	286	157	162	173	331	530	2,537
TOTAL											19,061

Table 8. CLSA Comprehensive HR2-Mail-outs by Health Registries Targeting Low-Education Areas: Number of Letters Sent by Age-Sex Groups

COMPREHENSIVE COHORT											
Province	Number of Mail-outs Conducted	Year	Number of People Mailed Out by Age-Sex Groups								TOTAL
			M 45-54	F 45-54	M 55-64	F 55-64	M 65-74	F 65-74	M 75-85	F 75-85	
NS	1	2014	695	439	531	506	1382	513	369	423	4,858
TOTAL											4,858

Table 9. Data Collection Sites

Province	Number of DCS's	DCS	Affiliated University/Research Institute
Alberta	1	Calgary Data Collection Site	University of Calgary
British Columbia	3	Victoria Data Collection Site, Vancouver Data Collection Site, Surrey Data Collection Site	University of Victoria (UVIC), University of British Columbia (UBC), Simon Fraser University (SFU)
Manitoba	1	Winnipeg Data Collection Site	University of Manitoba
Nova Scotia	1	Halifax Data Collection Site	Dalhousie University
Newfoundland and Labrador	1	St. John's Data Collection Site	Memorial University
Ontario	2	Hamilton Data Collection Site, Ottawa Data Collection Site	McMaster University, Bruyère Research Institute
Quebec	2	Montreal Data Collection Site, Sherbrooke Data Collection Site	McGill University, Research Institute of the McGill University Health Centre, Université de Sherbrooke

Table 10. Source of Participants for the CLSA Tracking by Sample and Province

Province	Sample			Total Number of participants in each province
	CCHS	HR	TS	
Alberta	352	0	1,755	2,107
British Columbia	425	86	2,109	2,620
Manitoba	350	470	664	1,484
New Brunswick	268	316	775	1,359
Newfoundland and Labrador	209	337	707	1,253
Nova Scotia	310	1,208	35	1,553
Ontario	746	612	3,364	4,722
Prince Edward Island	213	327	603	1,143
Quebec	736	0	2,872	3,608
Saskatchewan	314	454	624	1,392
Total Number of participants in each sample	**3,923**	**3,810**	**13,508**	**TOTAL=21,241**

Table 11. Source of Participants for the CLSA Comprehensive by Sample and Province

Province	Sample		Total Number of participants in each province
	HR	TS*	
Alberta	0	2,957	2,957
British Columbia	133	6,121	6,254
Manitoba	496	2,617	3,113
Newfoundland and Labrador	443	1,771	2,214
Nova Scotia	2,093	985	3,078
Ontario	964	5,454	6,418
Quebec	0	6,063	6,063
Total Number of participants in each sample	**4,129**	**25,968**	**TOTAL=30,097**

*** NuAge participants were treated as RDD participants.**

Table 12. CLSA Tracking Cohort Response Rates by Province and Sample and Overall

	AB	BC	MB	NB	NL	NS	ON	PE	QC	SK	CANADA
CCHS	**0.12**	**0.11**	**0.15**	**0.12**	**0.11**	**0.13**	**0.11**	**0.13**	**0.13**	**0.14**	**0.12**
RDD	0.09	0.11	0.10	0.13	0.09	-	0.10	0.13	0.15	0.09	0.11
RTS	0.01	0.01	0.01	0.01	0.01	0.02	0.01	-	0.02	0.01	0.01
TS	**0.07**	**0.10**	**0.09**	**0.10**	**0.08**	**0.02**	**0.09**	**0.13**	**0.13**	**0.07**	**0.10**
HR1	-	-	0.08	0.07	0.06	0.12	0.04	0.06	-	0.09	0.07
HR2	-	0.02	0.03	0.02	0.01	0.08	-	0.02	-	-	0.03
HR	**-**	**0.02**	**0.07**	**0.05**	**0.05**	**0.10**	**0.04**	**0.05**	**-**	**0.09**	**0.06**
OVERALL	**0.08**	**0.09**	**0.09**	**0.08**	**0.07**	**0.10**	**0.08**	**0.09**	**0.13**	**0.08**	**0.09**

Table 13. CLSA Comprehensive Cohort Response Rates by Province and Sample and Overall

	AB	BC	MB	NL	NS	ON	QC	CANADA
RDD	0.11	0.10	0.13	0.19	0.16	0.10	0.12	0.11
RTS	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.02
TS	**0.11**	**0.10**	**0.10**	**0.15**	**0.12**	**0.09**	**0.10**	**0.10**
HR1	-	0.02	0.09	0.06	0.16	0.09	-	0.09
HR2	-	-	-	-	0.08	-	-	0.08
HR	**-**	**0.02**	**0.09**	**0.06**	**0.14**	**0.09**	**-**	**0.09**
OVERALL	**0.11**	**0.09**	**0.10**	**0.12**	**0.13**	**0.09**	**0.10**	**0.10**

Table 14. CLSA Pooled Response Rates by province and Sample and Overall

	AB	BC	MB	NB	NL	NS	ON	PE	QC	SK	CANADA
CCHS	**0.12**	**0.11**	**0.15**	**0.12**	**0.11**	**0.13**	**0.11**	**0.13**	**0.13**	**0.14**	**0.12**
RDD	0.11	0.11	0.13	0.17	0.16	0.16	0.11	0.17	0.13	0.11	0.10
RTS	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-	0.02	0.01	0.02
TS	**0.10**	**0.11**	**0.11**	**0.12**	**0.13**	**0.11**	**0.10**	**0.17**	**0.12**	**0.08**	**0.09**
HR1	-	0.02	0.09	0.07	0.06	0.14	0.06	0.06	-	0.09	0.08
HR2	-	0.02	0.03	0.02	0.01	0.08	-	0.02	-	-	0.04
HR	**-**	**0.02**	**0.08**	**0.05**	**0.05**	**0.12**	**0.06**	**0.05**	**-**	**0.09**	**0.07**
OVERALL	**0.10**	**0.10**	**0.10**	**0.09**	**0.10**	**0.12**	**0.09**	**0.09**	**0.12**	**0.09**	**0.10**

Table 15. Total Weights Computed by Statistics Canada*

	Number of Eligible Participants Who Allowed Statistics Canada to Pass on Their CCHS Data to the CLSA	**Weights**		
		Minimum	**Maximum**	**Sum**
Survey+Contact	11,742	3.07	22,374.19	7,678,625.69
Survey Only	8,345	3.25	22,317.57	5,554,025.08
TOTAL	**20,087**			**13,232,650.77****

*** Statistics Canada re-calculated the CCHS weights for these 20,087 people to add to the total eligible population.**
**** The total in DCS areas was 3,746,315.46.**

Table 16. The Sampling Strata and the Number of Participants in Each in the Final CLSA Tracking Group

Stratum Number	Province	Sex	Age Group	DCS	Number of Participants	Stratum Number	Province	Sex	Age Group	DCS	Number of Participants
1	AB	female	45-54	DCS	48	**69**	NL	male	65-74	DCS	36
2	AB	female	45-54	Non-DCS	291	**70**	NL	male	65-74	Non-DCS	92
3	AB	female	55-64	DCS	65	**71**	NL	male	75-85	DCS	38
4	AB	female	55-64	Non-DCS	283	**72**	NL	male	75-85	Non-DCS	90
5	AB	female	65-74	DCS	44	**73**	NS	female	45-54	DCS	68
6	AB	female	65-74	Non-DCS	160	**74**	NS	female	45-54	Non-DCS	159
7	AB	female	75-85	DCS	28	**75**	NS	female	55-64	DCS	72
8	AB	female	75-85	Non-DCS	162	**76**	NS	female	55-64	Non-DCS	179
9	AB	male	45-54	DCS	36	**77**	NS	female	65-74	DCS	45
10	AB	male	45-54	Non-DCS	275	**78**	NS	female	65-74	Non-DCS	122
11	AB	male	55-64	DCS	48	**79**	NS	female	75-85	DCS	39
12	AB	male	55-64	Non-DCS	266	**80**	NS	female	75-85	Non-DCS	92
13	AB	male	65-74	DCS	38	**81**	NS	male	45-54	DCS	77
14	AB	male	65-74	Non-DCS	167	**82**	NS	male	45-54	Non-DCS	146
15	AB	male	75-85	DCS	17	**83**	NS	male	55-64	DCS	65
16	AB	male	75-85	Non-DCS	179	**84**	NS	male	55-64	Non-DCS	168
17	BC	female	45-54	DCS	137	**85**	NS	male	65-74	DCS	43
18	BC	female	45-54	Non-DCS	270	**86**	NS	male	65-74	Non-DCS	127
19	BC	female	55-64	DCS	129	**87**	NS	male	75-85	DCS	37
20	BC	female	55-64	Non-DCS	302	**88**	NS	male	75-85	Non-DCS	114
21	BC	female	65-74	DCS	72	**89**	ON	female	45-54	DCS	72
22	BC	female	65-74	Non-DCS	199	**90**	ON	female	45-54	Non-DCS	622
23	BC	female	75-85	DCS	77	**91**	ON	female	55-64	DCS	86
24	BC	female	75-85	Non-DCS	178	**92**	ON	female	55-64	Non-DCS	669
25	BC	male	45-54	DCS	110	**93**	ON	female	65-74	DCS	53
26	BC	male	45-54	Non-DCS	250	**94**	ON	female	65-74	Non-DCS	465
27	BC	male	55-64	DCS	135	**95**	ON	female	75-85	DCS	36
28	BC	male	55-64	Non-DCS	268	**96**	ON	female	75-85	Non-DCS	423
29	BC	male	65-74	DCS	76	**97**	ON	male	45-54	DCS	66
30	BC	male	65-74	Non-DCS	179	**98**	ON	male	45-54	Non-DCS	608
31	BC	male	75-85	DCS	88	**99**	ON	male	55-64	DCS	77
32	BC	male	75-85	Non-DCS	150	**100**	ON	male	55-64	Non-DCS	645
33	MB	female	45-54	DCS	120	**101**	ON	male	65-74	DCS	59
34	MB	female	45-54	Non-DCS	108	**102**	ON	male	65-74	Non-DCS	401
35	MB	female	55-64	DCS	129	**103**	ON	male	75-85	DCS	43
36	MB	female	55-64	Non-DCS	111	**104**	ON	male	75-85	Non-DCS	397
37	MB	female	65-74	DCS	57	**105**	PE	female	45-54	Non-DCS	165
38	MB	female	65-74	Non-DCS	84	**106**	PE	female	55-64	Non-DCS	165
39	MB	female	75-85	DCS	64	**107**	PE	female	65-74	Non-DCS	127
40	MB	female	75-85	Non-DCS	87	**108**	PE	female	75-85	Non-DCS	121
41	MB	male	45-54	DCS	97	**109**	PE	male	45-54	Non-DCS	160
42	MB	male	45-54	Non-DCS	127	**110**	PE	male	55-64	Non-DCS	151
43	MB	male	55-64	DCS	115	**111**	PE	male	65-74	Non-DCS	127
44	MB	male	55-64	Non-DCS	101	**112**	PE	male	75-85	Non-DCS	127
45	MB	male	65-74	DCS	63	**113**	QC	female	45-54	DCS	98
46	MB	male	65-74	Non-DCS	86	**114**	QC	female	45-54	Non-DCS	483
47	MB	male	75-85	DCS	62	**115**	QC	female	55-64	DCS	118
48	MB	male	75-85	Non-DCS	73	**116**	QC	female	55-64	Non-DCS	459
49	NB	female	45-54	Non-DCS	210	**117**	QC	female	65-74	DCS	72
50	NB	female	55-64	Non-DCS	212	**118**	QC	female	65-74	Non-DCS	277
51	NB	female	65-74	Non-DCS	138	**119**	QC	female	75-85	DCS	65
52	NB	female	75-85	Non-DCS	131	**120**	QC	female	75-85	Non-DCS	249
53	NB	male	45-54	Non-DCS	195	**121**	QC	male	45-54	DCS	80
54	NB	male	55-64	Non-DCS	201	**122**	QC	male	45-54	Non-DCS	446
55	NB	male	65-74	Non-DCS	143	**123**	QC	male	55-64	DCS	97
56	NB	male	75-85	Non-DCS	129	**124**	QC	male	55-64	Non-DCS	478
57	NL	female	45-54	DCS	59	**125**	QC	male	65-74	DCS	76
58	NL	female	45-54	Non-DCS	131	**126**	QC	male	65-74	Non-DCS	290
59	NL	female	55-64	DCS	64	**127**	QC	male	75-85	DCS	59
60	NL	female	55-64	Non-DCS	125	**128**	QC	male	75-85	Non-DCS	261
61	NL	female	65-74	DCS	38	**129**	SK	female	45-54	Non-DCS	217
62	NL	female	65-74	Non-DCS	88	**130**	SK	female	55-64	Non-DCS	221
63	NL	female	75-85	DCS	36	**131**	SK	female	65-74	Non-DCS	146
64	NL	female	75-85	Non-DCS	87	**132**	SK	female	75-85	Non-DCS	129
65	NL	male	45-54	DCS	45	**133**	SK	male	45-54	Non-DCS	189
66	NL	male	45-54	Non-DCS	128	**134**	SK	male	55-64	Non-DCS	215
67	NL	male	55-64	DCS	47	**135**	SK	male	65-74	Non-DCS	144
68	NL	male	55-64	Non-DCS	149	**136**	SK	male	75-85	Non-DCS	131

Table 17. The Sampling Strata and the Number of Participants in Each in the Final CLSA Comprehensive Group

Stratum Number	Province	Sex	Age Group	DCS	Number of Participants	Stratum Number	Province	Sex	Age Group	DCS	Number of Participants
1	AB	female	45-54	DCS	384	29	NL	male	45-54	DCS	274
2	AB	female	55-64	DCS	509	30	NL	male	55-64	DCS	343
3	AB	female	65-74	DCS	371	31	NL	male	65-74	DCS	264
4	AB	female	75-85	DCS	253	32	NL	male	75-85	DCS	201
5	AB	male	45-54	DCS	329	33	NS	female	45-54	DCS	391
6	AB	male	55-64	DCS	492	34	NS	female	55-64	DCS	499
7	AB	male	65-74	DCS	375	**35**	NS	female	65-74	DCS	389
8	AB	male	75-85	DCS	244	**36**	NS	female	75-85	DCS	270
9	BC	female	45-54	DCS	831	**37**	NS	male	45-54	DCS	378
10	BC	female	55-64	DCS	1030	**38**	NS	male	55-64	DCS	460
11	BC	female	65-74	DCS	724	**39**	NS	male	65-74	DCS	424
12	BC	female	75-85	DCS	573	**40**	NS	male	75-85	DCS	267
13	BC	male	45-54	DCS	782	**41**	ON	female	45-54	DCS	803
14	BC	male	55-64	DCS	980	**42**	ON	female	55-64	DCS	1070
15	BC	male	65-74	DCS	737	**43**	ON	female	65-74	DCS	780
16	BC	male	75-85	DCS	597	**44**	ON	female	75-85	DCS	554
17	MB	female	45-54	DCS	415	**45**	ON	male	45-54	DCS	781
18	MB	female	55-64	DCS	527	**46**	ON	male	55-64	DCS	1051
19	MB	female	65-74	DCS	373	**47**	ON	male	65-74	DCS	788
20	MB	female	75-85	DCS	279	**48**	ON	male	75-85	DCS	591
21	MB	male	45-54	DCS	366	**49**	QC	female	45-54	DCS	792
22	MB	male	55-64	DCS	511	**50**	QC	female	55-64	DCS	1075
23	MB	male	65-74	DCS	367	**51**	QC	female	65-74	DCS	786
24	MB	male	75-85	DCS	275	**52**	QC	female	75-85	DCS	510
25	NL	female	45-54	DCS	309	**53**	QC	male	45-54	DCS	760
26	NL	female	55-64	DCS	379	**54**	QC	male	55-64	DCS	930
27	NL	female	65-74	DCS	265	**55**	QC	male	65-74	DCS	719
28	NL	female	75-85	DCS	179	**56**	QC	male	75-85	DCS	491

Table 18. NHS 2011 (Weighted) Totals within Each Province Crossed with Individual DCS (and Provincial Non-DCS) Crossed with Individual Education by Sex Crossed with Age Groups

			SEX																					
			Female										Male											
			Age group										Age group											
			45-48		49-54		55-64		65-74		75-85		45-48		49-54		55-64		65-74		75-85			
PR	DCS	Individual Education (HCDD_7V)	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	OVERALL	%
AB	Calgary DCS	Higher Education Lower	10,765	29.9	16,520	30.8	18,925	30.5	7,950	25.5	4,035	20.2	8,465	23.5	12,670	24.1	13,790	22.0	5,650	20.0	2,470	15.6	101,240	25.4
		Higher Education Upper	11,805	32.8	15,165	28.2	15,555	25.1	4,910	15.7	1,945	9.7	12,640	35.0	16,235	30.9	21,000	33.5	8,110	28.7	3,385	21.4	110,750	27.8
		Low Education	3,205	8.9	5,320	9.9	7,390	11.9	7,585	24.3	7,120	35.6	3,790	10.5	6,065	11.5	6,650	10.6	4,825	17.1	4,375	27.6	56,325	14.1
		Medium Education	10,200	28.4	16,705	31.1	20,090	32.4	10,785	34.5	6,900	34.5	11,200	31.0	17,570	33.4	21,230	33.9	9,660	34.2	5,610	35.4	129,950	32.6
	NonDCS	Higher Education Lower	23,525	32.5	33,745	30.5	39,910	28.5	19,325	24.6	8,580	18.3	16,455	23.0	23,200	20.8	28,805	20.6	12,390	16.8	5,090	12.2	211,025	23.8
		Higher Education Upper	14,310	19.7	18,010	16.3	24,710	17.6	8,640	11.0	2,835	6.0	12,550	17.6	16,195	14.5	25,550	18.2	12,090	16.4	4,440	10.7	139,330	15.7
		Low Education	9,300	12.8	15,665	14.2	21,890	15.6	22,780	29.0	20,745	44.1	10,520	14.7	19,060	17.1	22,470	16.0	20,250	27.5	16,610	39.9	179,290	20.2
		Medium Education	25,360	35.0	43,240	39.1	53,580	38.2	27,915	35.5	14,850	31.6	31,960	44.7	53,050	47.6	63,330	45.2	28,855	39.2	15,480	37.2	357,620	40.3
BC	SFU_UBC DCS	Higher Education Lower	23,210	30.6	32,500	29.5	38,135	26.2	19,920	23.6	9,440	17.3	16,815	23.4	24,275	23.7	30,010	21.7	14,785	19.5	7,275	15.9	216,365	23.9
		Higher Education Upper	23,990	31.6	29,805	27.1	35,420	24.4	14,090	16.7	4,950	9.1	24,515	34.1	30,340	29.6	41,855	30.3	19,655	26.0	8,970	19.6	233,590	25.8
		Low Education	5,955	7.9	10,215	9.3	19,970	13.7	21,325	25.2	21,075	38.6	6,375	8.9	10,965	10.7	16,620	12.0	13,980	18.5	13,250	29.0	139,730	15.5
		Medium Education	22,680	29.9	37,490	34.1	51,930	35.7	29,140	34.5	19,200	35.1	24,175	33.6	36,990	36.1	49,645	35.9	27,210	36.0	16,230	35.5	314,690	34.8
	Victoria DCS	Higher Education Lower	3,765	32.4	5,940	30.5	8,985	30.0	5,150	29.3	3,285	25.4	2,725	24.6	3,670	22.6	6,120	22.4	3,300	19.9	1,835	18.6	44,775	25.9
		Higher Education Upper	3,530	30.4	5,230	26.9	8,800	29.3	3,935	22.4	1,805	14.0	2,910	26.3	4,235	26.0	8,420	30.8	5,480	33.1	2,645	26.8	46,990	27.2
		Low Education	870	7.5	1,235	6.3	1,980	6.6	2,445	13.9	3,305	25.6	765	6.9	1,535	9.4	2,240	8.2	2,170	13.1	1,830	18.5	18,375	10.6
		Medium Education	3,455	29.7	7,055	36.3	10,220	34.1	6,075	34.5	4,515	35.0	4,680	42.2	6,830	42.0	10,520	38.5	5,600	33.8	3,560	36.1	62,510	36.2
	NonDCS	Higher Education Lower	18,070	34.9	27,510	33.0	39,005	30.3	21,415	25.9	10,880	21.8	10,835	22.7	17,210	21.3	25,475	20.6	14,240	17.6	7,160	15.1	191,800	24.7
		Higher Education Upper	8,325	16.1	12,080	14.5	20,685	16.0	8,940	10.8	3,050	6.1	6,855	14.3	10,130	12.5	20,465	16.5	13,475	16.7	5,340	11.3	109,345	14.1
		Low Education	5,270	10.2	9,605	11.5	16,980	13.2	20,935	25.3	17,875	35.7	7,135	14.9	13,605	16.8	19,815	16.0	18,230	22.6	15,385	32.5	144,835	18.6
		Medium Education	20,100	38.8	34,045	40.9	52,215	40.5	31,475	38.0	18,200	36.4	22,960	48.0	39,835	49.3	58,190	46.9	34,870	43.1	19,445	41.1	331,335	42.6
MB	Manitoba DCS	Higher Education Lower	6,330	29.5	9,150	28.6	11,385	25.0	5,450	21.4	3,160	16.3	5,045	23.9	6,930	21.9	8,750	21.4	4,095	17.7	1,830	13.4	62,125	22.7
		Higher Education Upper	5,805	27.1	7,640	23.9	10,025	22.0	3,655	14.3	1,450	7.5	5,550	26.3	7,435	23.5	10,160	24.8	4,745	20.5	2,265	16.6	58,730	21.4
		Low Education	2,325	10.8	4,155	13.0	7,585	16.7	6,950	27.2	7,970	41.1	3,055	14.5	5,435	17.2	6,495	15.9	5,800	25.0	4,365	32.0	54,135	19.7
		Medium Education	6,990	32.6	11,015	34.5	16,480	36.2	9,465	37.1	6,810	35.1	7,430	35.2	11,825	37.4	15,490	37.9	8,530	36.8	5,190	38.0	99,225	36.2
	NonDCS	Higher Education Lower	3,850	29.7	5,830	30.1	7,285	26.5	3,970	22.4	2,265	18.7	2,640	21.9	3,840	19.8	4,660	16.5	2,600	14.5	940	8.9	37,880	21.3
		Higher Education Upper	1,890	14.6	2,295	11.8	3,890	14.2	1,410	8.0	415	3.4	1,385	11.5	1,810	9.3	3,835	13.6	2,210	12.3	830	7.8	19,970	11.2
		Low Education	2,070	16.0	3,555	18.3	6,100	22.2	6,490	36.6	6,325	52.3	2,765	22.9	5,335	27.5	8,315	29.4	7,185	39.9	5,525	52.2	53,665	30.2
		Medium Education	5,145	39.7	7,695	39.7	10,200	37.1	5,840	33.0	3,090	25.5	5,290	43.8	8,385	43.3	11,490	40.6	5,990	33.3	3,285	31.0	66,410	37.3
NB	NonDCS	Higher Education Lower	7,400	29.9	10,055	26.3	12,460	22.0	7,365	21.3	3,465	16.1	5,305	22.6	6,780	19.1	9,010	16.4	4,445	13.9	1,585	9.5	67,870	20.1
		Higher Education Upper	4,630	18.7	5,305	13.9	8,255	14.5	3,145	9.1	1,060	4.9	3,645	15.5	4,785	13.5	8,010	14.6	4,115	12.8	1,480	8.9	44,430	13.1
		Low Education	3,170	12.8	7,040	18.4	14,145	24.9	13,055	37.8	11,475	53.4	4,220	18.0	8,025	22.6	14,145	25.8	12,435	38.8	8,235	49.5	95,945	28.4
		Medium Education	9,510	38.5	15,870	41.5	21,890	38.6	10,960	31.7	5,480	25.5	10,305	43.9	15,980	44.9	23,675	43.2	11,045	34.5	5,340	32.1	130,055	38.4

Table 18. Continued

			SEX																					
			femaLow Education										maLow Education											
			Age group										Age group											
		Individual Education	45-48		49-54		55-64		65-74		75-85		45-48		49-54		55-64		65-74		75-85			
PR	DCS	(HCDD_7V)	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	OVERALL	%
NL	Memorial DCS	Higher Education Lower	2,390	37.9	3,450	35.8	4,040	28.5	1,700	22.8	705	16.8	1,760	29.1	2,310	26.3	2,690	22.8	1,425	19.9	505	18.8	20,975	26.8
		Higher Education Upper	1,385	22.0	1,735	18.0	2,735	19.3	1,085	14.5	210	5.0	1,300	21.5	1,575	17.9	2,575	21.8	1,455	20.4	315	11.7	14,370	18.4
		Low Education	500	7.9	1,195	12.4	2,240	15.8	2,040	27.3	1,925	45.9	760	12.6	1,345	15.3	1,965	16.6	1,590	22.3	890	33.1	14,450	18.5
		Medium Education	2,030	32.2	3,250	33.7	5,170	36.4	2,635	35.3	1,350	32.2	2,230	36.9	3,560	40.5	4,585	38.8	2,675	37.4	980	36.4	28,465	36.4
	NonDCS	Higher Education Lower	3,005	29.1	4,705	27.8	5,510	20.0	2,145	12.8	790	8.5	2,340	22.5	3,265	20.4	4,690	17.3	2,175	13.2	485	6.3	29,110	18.4
		Higher Education Upper	990	9.6	1,065	6.3	2,000	7.3	600	3.6	105	1.1	800	7.7	1,020	6.4	2,255	8.3	1,105	6.7	150	1.9	10,090	6.4
		Low Education	2,405	23.3	4,570	27.0	10,210	37.0	9,420	56.3	6,510	70.3	2,945	28.3	4,765	29.8	9,140	33.6	7,390	44.8	5,230	68.0	62,585	39.5
		Medium Education	3,940	38.1	6,580	38.9	9,850	35.7	4,575	27.3	1,860	20.1	4,325	41.5	6,960	43.5	11,080	40.8	5,820	35.3	1,830	23.8	56,820	35.8
NS	Dalhousie DCS	Higher Education Lower	4,035	33.0	5,615	30.7	6,915	28.2	3,625	27.3	1,945	22.9	2,615	22.7	4,005	23.9	3,970	18.8	2,200	18.5	1,070	17.8	35,995	25.0
		Higher Education Upper	3,825	31.3	4,825	26.3	6,115	24.9	2,035	15.3	985	11.6	3,210	27.9	4,115	24.6	6,250	29.6	2,975	25.1	1,120	18.6	35,455	24.6
		Low Education	1,080	8.8	1,865	10.2	3,555	14.5	3,090	23.2	2,935	34.5	1,335	11.6	2,205	13.2	2,880	13.6	2,415	20.4	1,550	25.8	22,910	15.9
		Medium Education	3,280	26.8	6,010	32.8	7,955	32.4	4,550	34.2	2,635	31.0	4,355	37.8	6,420	38.3	8,010	37.9	4,275	36.0	2,270	37.8	49,760	34.5
	NonDCS	Higher Education Lower	6,495	35.4	9,700	32.9	12,830	28.3	7,275	24.3	3,780	20.8	4,215	24.8	6,050	21.9	8,710	19.8	3,785	13.1	1,500	10.4	64,340	23.5
		Higher Education Upper	2,945	16.0	4,720	16.0	5,970	13.2	2,555	8.5	815	4.5	1,840	10.8	3,025	10.9	6,730	15.3	3,610	12.5	1,405	9.7	33,615	12.3
		Low Education	2,490	13.6	5,010	17.0	10,055	22.2	11,245	37.6	9,025	49.7	3,700	21.7	6,615	23.9	10,805	24.5	10,390	36.0	6,925	47.8	76,260	27.9
		Medium Education	6,430	35.0	10,085	34.2	16,500	36.4	8,860	29.6	4,545	25.0	7,270	42.7	11,960	43.3	17,815	40.4	11,110	38.4	4,650	32.1	99,225	36.3
ON	Hamilton DCS	Higher Education Lower	9,215	36.2	12,325	33.6	14,530	29.2	7,300	23.1	4,190	17.6	7,290	29.4	9,560	27.1	11,120	24.5	5,520	19.1	2,700	15.0	83,750	26.2
		Higher Education Upper	6,730	26.4	7,915	21.6	9,560	19.2	3,640	11.5	1,255	5.3	6,760	27.3	7,995	22.7	11,040	24.3	6,045	20.9	2,475	13.8	63,415	19.8
		Low Education	2,105	8.3	4,145	11.3	7,570	15.2	9,435	29.9	10,300	43.2	2,315	9.3	4,250	12.0	6,715	14.8	7,385	25.6	6,345	35.3	60,565	18.9
		Medium Education	7,430	29.2	12,330	33.6	18,065	36.3	11,225	35.5	8,105	34.0	8,425	34.0	13,475	38.2	16,515	36.4	9,945	34.4	6,440	35.9	111,955	35.0
	Ottawa DCS	Higher Education Lower	7,775	29.2	11,415	29.4	13,325	26.5	7,420	24.6	3,975	20.6	6,425	24.8	8,800	24.9	10,235	21.8	4,665	17.9	2,475	16.3	76,510	24.3
		Higher Education Upper	10,560	39.7	13,375	34.4	16,835	33.5	7,405	24.5	2,605	13.5	11,425	44.2	13,865	39.3	19,845	42.3	11,080	42.5	4,695	30.9	111,690	35.5
		Low Education	1,555	5.8	3,035	7.8	4,865	9.7	5,975	19.8	5,790	30.1	2,065	8.0	3,125	8.9	4,280	9.1	3,100	11.9	3,340	22.0	37,130	11.8
		Medium Education	6,715	25.2	11,050	28.4	15,175	30.2	9,415	31.2	6,895	35.8	5,960	23.0	9,500	26.9	12,575	26.8	7,220	27.7	4,700	30.9	89,205	28.4
	NonDCS	Higher Education Lower	123,290	32.9	166,325	30.8	195,415	26.8	97,435	21.5	48,885	16.1	92,135	26.1	126,550	24.7	151,105	22.2	71,945	17.4	33,965	14.0	1,107,050	24.0
		Higher Education Upper	96,155	25.6	118,820	22.0	143,935	19.7	58,145	12.8	22,440	7.4	90,090	25.5	119,235	23.3	160,350	23.5	83,030	20.0	36,095	14.9	928,295	20.2
		Low Education	37,250	9.9	64,245	11.9	124,795	17.1	145,755	32.1	135,495	44.7	41,460	11.8	73,125	14.3	110,270	16.2	108,720	26.2	89,930	37.1	931,045	20.2
		Medium Education	118,510	31.6	189,790	35.2	266,340	36.5	152,450	33.6	96,640	31.8	129,040	36.6	193,345	37.7	259,300	38.1	150,630	36.4	82,160	33.9	1,638,205	35.6
PE	NonDCS	Higher Education Lower	1,470	33.2	2,570	36.3	3,420	33.0	1,955	30.6	1,000	25.2	915	21.1	1,365	21.6	1,720	17.1	685	11.8	445	13.6	15,545	25.1
		Higher Education Upper	980	22.1	1,045	14.8	1,945	18.8	570	8.9	215	5.4	835	19.3	975	15.5	1,785	17.8	860	14.9	360	11.0	9,570	15.4
		Low Education	445	10.0	690	9.7	1,630	15.7	1,840	28.8	1,760	44.3	795	18.4	1,540	24.4	2,235	22.3	1,710	29.5	1,415	43.2	14,060	22.7
		Medium Education	1,535	34.7	2,775	39.2	3,355	32.4	2,025	31.7	1,000	25.2	1,785	41.2	2,425	38.5	4,295	42.8	2,535	43.8	1,055	32.2	22,785	36.8

Table 18. Continued

			SEX																					
			femaLow Education										maLow Education											
			Age group										Age group											
			45-48		49-54		55-64		65-74		75-85		45-48		49-54		55-64		65-74		75-85			
PR	DCS	Individual Education (HCDD_7V)	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	Sum	%	OVERALL	%
QC	McGill DCS	Higher Education Lower	22,215	**26.8**	30,395	**24.6**	37,910	**21.9**	21,815	**18.5**	10,105	**12.2**	18,315	**22.2**	25,740	**21.4**	29,845	**19.0**	14,885	**15.2**	6,740	**11.4**	217,965	**19.9**
		Higher Education Upper	24,620	**29.7**	31,695	**25.7**	38,815	**22.4**	17,330	**14.7**	6,125	**7.4**	25,760	**31.2**	33,685	**28.0**	42,940	**27.3**	22,245	**22.7**	9,525	**16.1**	252,740	**23.0**
		Low Education	9,695	**11.7**	17,460	**14.2**	33,030	**19.0**	39,835	**33.9**	41,215	**49.8**	10,610	**12.8**	18,750	**15.6**	26,850	**17.1**	25,645	**26.2**	22,715	**38.4**	245,805	**22.4**
		Medium Education	26,270	**31.7**	43,795	**35.5**	63,695	**36.7**	38,630	**32.8**	25,360	**30.6**	27,890	**33.8**	42,000	**34.9**	57,435	**36.6**	35,025	**35.8**	20,140	**34.1**	380,240	**34.7**
	Sherbrooke DCS	Higher Education Lower	1,680	**23.6**	2,895	**24.4**	4,095	**22.1**	2,280	**19.2**	995	**13.5**	1,405	**19.7**	1,855	**17.0**	3,010	**17.2**	1,335	**11.5**	525	**9.9**	20,075	**18.4**
		Higher Education Upper	1,390	**19.5**	1,905	**16.0**	2,835	**15.3**	1,470	**12.4**	570	**7.8**	1,325	**18.6**	1,705	**15.6**	3,055	**17.5**	2,135	**18.4**	765	**14.4**	17,155	**15.7**
		Low Education	985	**13.8**	2,095	**17.6**	3,785	**20.4**	3,950	**33.3**	3,525	**48.0**	1,120	**15.7**	2,390	**21.9**	3,635	**20.8**	3,295	**28.4**	2,290	**43.2**	27,070	**24.8**
		Medium Education	3,075	**43.1**	4,980	**41.9**	7,850	**42.3**	4,160	**35.1**	2,255	**30.7**	3,290	**46.1**	4,965	**45.5**	7,800	**44.6**	4,835	**41.7**	1,720	**32.5**	44,930	**41.1**
	NonDCS	Higher Education Lower	43,840	**28.2**	61,985	**24.6**	80,420	**22.5**	41,365	**18.6**	14,485	**11.5**	33,815	**22.2**	46,020	**18.7**	62,645	**17.9**	29,465	**13.8**	9,060	**9.3**	423,100	**19.5**
		Higher Education Upper	25,955	**16.7**	33,780	**13.4**	42,810	**12.0**	18,285	**8.2**	4,515	**3.6**	23,660	**15.5**	33,515	**13.6**	51,980	**14.9**	29,925	**14.0**	9,160	**9.4**	273,585	**12.6**
		Low Education	20,360	**13.1**	44,125	**17.5**	76,220	**21.4**	85,465	**38.4**	72,025	**57.2**	23,785	**15.6**	50,995	**20.7**	73,005	**20.9**	69,195	**32.3**	47,605	**48.6**	562,780	**25.9**
		Medium Education	65,270	**42.0**	111,820	**44.4**	157,415	**44.1**	77,270	**34.7**	34,820	**27.7**	71,010	**46.6**	115,875	**47.0**	161,870	**46.3**	85,435	**39.9**	32,040	**32.7**	912,825	**42.0**
SK	NonDCS	Higher Education Lower	9,435	**33.0**	13,210	**29.4**	17,340	**28.5**	9,395	**25.3**	5,875	**19.7**	4,790	**17.9**	7,730	**17.4**	9,795	**16.1**	4,250	**12.2**	2,390	**10.4**	84,210	**21.5**
		Higher Education Upper	5,240	**18.3**	7,250	**16.1**	9,560	**15.7**	3,360	**9.0**	1,385	**4.7**	4,395	**16.4**	5,565	**12.6**	9,920	**16.3**	5,550	**15.9**	1,855	**8.1**	54,080	**13.8**
		Low Education	3,265	**11.4**	5,845	**13.0**	10,355	**17.0**	12,105	**32.6**	13,585	**45.6**	4,890	**18.3**	8,850	**20.0**	11,925	**19.6**	12,140	**34.8**	11,375	**49.4**	94,335	**24.1**
		Medium Education	10,675	**37.3**	18,595	**41.4**	23,655	**38.8**	12,320	**33.1**	8,920	**30.0**	12,700	**47.4**	22,180	**50.0**	29,265	**48.1**	12,935	**37.1**	7,410	**32.2**	158,655	**40.5**

Table 19.1 Number of Postal Codes per Province by % of People with Lower Levels of Education*

	≥ 40		≥ 45		≥ 50		≥ 55		≥ 60		≥ 65		TOTAL
Province	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**
AB	30648	38.2	21416	26.7	13961	17.4	7685	9.6	**3494**	4.4	1557	1.9	80235
BC	43558	38.3	28648	25.2	17823	15.7	8743	7.7	**3944**	3.5	1703	1.5	113732
MB	12984	53.7	9868	40.8	7097	29.3	3981	16.5	2353	9.7	**1314**	5.4	24197
NB	41711	70.2	33587	56.5	25667	43.2	17477	29.4	**11157**	18.8	6197	10.4	59407
NL	3705	34.1	2303	21.2	1524	14.0	862	7.9	**553**	5.1	264	2.4	10872
NS	12356	43.9	8277	29.4	5537	19.7	2759	9.8	**1461**	5.2	464	1.7	28117
ON	120593	43.7	85631	31.0	60167	21.8	35175	12.7	18861	6.8	**8862**	3.2	276083
PE	1077	32.6	681	20.6	374	11.3	**243**	7.3	210	6.4	137	4.1	3307
QC	86713	41.0	57404	27.2	37366	17.7	19204	9.1	9208	4.4	**3840**	1.8	211384
SK	11932	54.0	9049	41.0	7001	31.7	4204	19.0	**2245**	10.2	914	4.1	22087
TOTAL	**365277**	**44.0**	**256864**	**31.0**	**176517**	**21.3**	**100333**	**12.1**	**53486**	**6.4**	**25252**	**3.0**	**829421**

*** Percentage of people aged 25-64 and 65+ in DAs (See Appendix 1)**

Table 19.2 Number of Postal Codes per Province

Province	**Cut Point**	**Number of Postal Codes**	**%**
AB	60	3494	4.4
BC	60	3944	3.5
MB	65	1314	5.4
NB	60	11157	18.8
NL	60	553	5.1

Province	**Cut Point**	**Number of Postal Codes**	**%**
NS	60	1461	5.2
ON	65	8862	3.2
PE	55	243	7.3
QC	65	3840	1.8
SK	60	2245	10.2

APPENDICES

Appendix 1. Determination of the 'Low-Education' Dissemination Areas

As early indications showed that the proportion of participants with lower levels of education (defined as having no education beyond high school) was below that in the population, it was decided to try to oversample people with lower levels of education. The first step was to estimate how many people had to be sampled in each province to achieve the desired number of participants with high school or lower education.

The next step was to identify 'low education (LowEd)' areas – areas in which the proportion of people with high school or less education was relatively high. The CLSA used data from the 2006 Census. These data were the latest available that provided the required information. The aim was to restrict sampling to those areas, so that the 'yield' of people with lower levels of education would be relatively high, and the sampling process would thus be more efficient.

For confidentiality reasons, the CLSA did not have access to individual level data. Rather the data were grouped into age groups 25-64 and 65+, and according to 'Dissemination areas (DAs)' – see footnote 5. The CLSA identified DAs in which the proportion of people with high school or lower education in both age groups exceeded various percentages. The number of postal codes (PCs) in those DAs was determined (see Table 19.1). For each province, an estimate was made of the number of people who had to be sampled to achieve the target number of participants.

Using this figure and Table 19.1, an appropriate cut-point was chosen for the proportion of people with 'LowEd'. The cut-point was specific to the individual province and was intended to balance the efficiency of sampling with the need to obtain the target number of participants. As well, an average of no more than one person in any age-sex group was sampled per PC to ensure that the average total number of people sampled per PC was not so high that the sampling would be too concentrated in these PCs. The DAs within each province were divided into those within the catchment area of the Data Collection Sites, and those outside that area. Sampling for CLSA Tracking was then restricted to areas outside the catchment areas. In some provinces – PE, NB, SK – there was no DCS, so this included the whole province.

For example, Table 19.2 shows that in SK, there were 2,245 Postal Codes covered by the Dissemination areas in which at least 60% of people had 'low education'. Targeted sampling for the 'LowEd' mail-out (HR2) and for the second phase of telephone sampling (RTS) was restricted to these postal codes.

Appendix 2. CCHS Response Rate Calculation

CCHS response rates by province and age groups (A_h) were provided by Statistics Canada. The same rates were used for the male and female people in the same age groups within province.

Within province:

Proportion of people who provided their contact information to the CLSA in sex*age group h (B_h): (Number of pre-recruits in sex*age group h)/Number of people aged 45-85 from CCHS in sex*age group h.

Recruitment Rate in sex*age group h (C_h): Number of recruits in sex*age group h/(Number of pre-recruits in sex*age group h - Number of ineligible people because of their proxy mode and language in sex*age group h - Number of people contacted by the CLSA team and found ineligible in sex*age group h - Number of people not contacted by the CLSA team and estimated ineligible in sex*age group h).

Participation Rate in sex*age group h (D_h): Number of participants in sex*age group h/(Number of recruits in sex*age group h - Number of people unreachable and estimated ineligible in sex*age group h - Number of pre-recruits contacted but found ineligible in sex*age group h).

Response Rate 1= $B_h \times C_h \times D_h$

Response Rate 2=$A_h \times B_h \times C_h \times D_h$

Overall rates for provinces and Canada can be calculated by modifying the same formulas above.

Appendix 3. Provincial Health Registry Mail-outs Response Rate Calculation

1. Rates for each age and sex group within province

Pre-Recruitment Rate (A_h)= $r_h/(n_h - x_h)$,

where n_h is number of people sampled and sent letters in each age and sex group within province, r_h is number of people who replied to the CLSA in each age and sex group within province, x_h is number of returned mail-outs in each age and sex group within province.

Contact Rate (B_h)= s_h/r_h

where s_h is number of people contacted by the CLSA in each age and sex group within province

Full Participation Rate (C_h)=$t_h/(s_h - y_h)$

where t_h is number of peple who both completed baseline interview and provided consent form (participants) in each age and sex group within province, y_h is number of other ineligible in each age and sex group within province.

Response Rate for each age and sex group within province=$A_h \times B_h \times C_h$

2. Overall Response Rates by Provinces:

Pre-Recruitment Rate (A)= $r/(n - x)$

Contact Rate (B)= s/r

Full Participation Rate (C)= $t/(s - y)$

Response Rate=$A \times B \times C$

Appendix 4. Telephone Sampling Response Rate Calculation

1. N: Numbers of TNs in ASDE data set
2. n: Numbers of TNs called by Leger/CLSA CATI
3. $P(TNs\ called) = P_c = \frac{n}{N}$
4. $P(TNs\ valid|called) = P_{val} = \frac{n_{val}}{n}$
5. $P(TNs\ answered|valid) = P_{ans} = \frac{n_{ans}}{n_{val}}$
6. $P(TNs\ residence|answered) = P_{res} = \frac{n_{res}}{n_{ans}}$
7. $P(age\ eligible\ established|residence) = P_{age.est} = \frac{n_{age.est}}{n_{res}}$
8. $P(age\ eligible\ HHs|age\ eligible\ established) = P_{age} = \frac{n_{age}}{n_{age.est}}$
9. $P(rostered\ HHs|age\ eligible\ HHs) = P_{rost} = \frac{n_{rost}}{n_{age}}$
10. $'P'(individual\ selection) = P_{ind} = \frac{1}{k}$
    $k =$ the number of age-eligible people in contacted and co-operating household

10'. $P(selected\ people|rostered\ HHs) = P_{sel} = \frac{\#\ of\ people\ selected}{\#\ of\ people\ rostered*} = \frac{n_{sel}}{n_{rost}}$

11. $P(individual\ agrees|selected\ people) = P_{agr} = \frac{\#\ of\ people\ agreeing}{\#\ of\ people\ selected} = \frac{n_{agr}}{n_{sel}}$
12. $P(full\ eligible|individual\ agrees) = P_{fe} = \frac{n_{fe}}{n_{agr}}$
13. $P(coord|agreeing, full\ eligible) = P_{coord} = \frac{n_{coord}}{n_{fe}}$ Provides contact information ('co-ordinates')
14. $P(CLSA\ contact|coord) = Contact\ Rate = P_{CLSA} = \frac{n_{CLSA}}{n_{coord}}$
15. $P(agree\ to\ participate|CLSA\ contact) = P_{agree} = \frac{n_{agree}}{n_{CLSA}}$
16. $P(interview|agree\ to\ participate) = P_{int} = \frac{n_{int}}{n_{agree}}$
17. $P(consent\ form\ returned|interview) = P_{cons} = \frac{n_{cons}}{n_{int}}$
18. $P(full\ participant|prerecruit) = Full\ Participation\ Rate = P_{fullpart} = \quad P_{agree} \times P_{int} \times P_{cons} = \frac{n_{cons}}{n_{coord}}$

Estimated number of eligible people=$N \times P_{val} \times P_{res} \times P_{age} \times \frac{n_{age/sex}}{n_{rost}} \times P_{fe}$

Estimated number of eligible HHs=$N \times P_{val} \times P_{res} \times P_{age}$

HH Enumeration Rate=$P_{ans} \times P_{age.est} \times P_{rost}$

Pre-Recruitment Rate= HH Enumeration Rate$\times P_{agr} \times P_{coord}$

Conversion Rate=P_{conv}=$P_{CLSA} \times P_{fullpart}$

Response Rate=Pre-Recruitment Rate$\times P_{conv}$= $P_{ans} \times P_{age.est} \times P_{rost} \times P_{agr} \times P_{coord} \times P_{conv}$

Note: Multiplication of probabilities assumes Markov dependence of the separate components. To the extent this assumption is wrong, the final rates are not necessarily interpretable as probabilities.

* Number of people in the rostered households.

REFERENCES

Statistics Canada 2011. Census Dictionary, Census Year 2011. Catalogue no. 98-301-X2011001.

Statistics Canada [2010]. Residential Telephone Service Survey (RTSS)

Raina P, et al. Canadian Longitudinal Study on Aging (CLSA) Protocol. https://clsa-elcv.ca/doc/511. Accessed 03.11.2016.

Quebec Longitudinal Study on Nutrition and Aging (NuAge), http://www.rqrv.com/en/init_NuAge.php, Accessed 03.11.2016.

National Household Survey User Guide, https://www12.statcan.gc.ca/nhs-enm/2011/ref/nhs-enm_guide/index-eng.cfm, Accessed 03.09.2020