

The Canadian Community Health Survey: Building on the Success from the Past

Yves Béland, Vincent Dale, Johane Dufour, and Marc Hamel
 Statistics Canada, R.H. Coats Bldg, Ottawa (ON), Canada, K1A 0T6 (yves.beland@statcan.ca)

Key words: cross-sectional, regional survey, mixed-mode data collection.

1. Introduction

The need for health information, at all levels, is more critical than ever before. Policy makers and health care professionals are concerned with the increased demand anticipated for health care services due to the ageing of the population. With continuing advancements in the effectiveness of treatments, balancing costs and potential enhanced health outcomes will be a challenge. Health-care planners need basic surveillance information on health status and upstream determinants of health to better anticipate demands and evaluate services.

In the spring of 1999, Statistics Canada received funding through the *Health Information Roadmap Initiative* to develop and implement a new national health survey. The Canadian Community Health Survey (CCHS) was designed to address priority health determinants and health outcome data gaps and to help modernize health information in Canada (CIHI; 1999a, 1999b). CCHS, for which data collection began in 2000, consists of two cross-sectional surveys conducted over a two-year repeating cycle (Béland, Bailie, Catlin and Singh, 2000). The first survey, referred as cycle x.1, is designed to collect data from a sample of more than 125,000 respondents to provide information to Regional Health Authorities on a wide variety of topics such as health status, health care utilization and determinants of health. The second survey (cycle x.2) is designed to collect data from a smaller sample (approximately 30,000 respondents) to provide information to provinces on a specific health topic (mental health, nutrition, etc.).

The number of stakeholders, as well as the requirements for population health information, has increased since the inception of the CCHS. In keeping with its short history of being a very flexible survey program, important changes to the CCHS are being contemplated to respond to the increasing demand for information. The three areas under review are the survey design, the design of the questionnaire, and the dissemination strategy

This paper describes the main elements of the redesigned CCHS program. Section 2 gives an

overview of the first six years of the CCHS and how the survey program has adapted to meet the various requirements. Section 3 provides details on the steps that led to the survey redesign. Section 4 describes the main elements of a more responsive CCHS program. Details on the sample design of the CCHS are provided in section 5. Finally, areas for future developments of the Statistics Canada Cross-Sectional Health Survey Program are given in Section 6.

2. The First Six Years of CCHS

The CCHS was launched in September 2000 following extensive consultations with various levels of stakeholders across Canada, including Health Canada, provincial and territorial ministries of health, health regions, and various academics and non-governmental organizations involved in population health. The various elements of the initial design of the survey were developed to respond to the priority needs of stakeholders (Béland *et al.*, 2000). As the list of stakeholders grew in time, the survey elements were adapted to better respond to increased needs.

2.1. Questionnaire

The questionnaire for the CCHS is designed to be very flexible. Within the 45-minute interview, all health regions (more than 120 regions) are allowed to select 10 minutes of optional content, at their discretion, from a predetermined list of content modules. The rest of the interview consists of common content, which is asked of all respondents and based on the highest priorities of the main stakeholders. Most of the CCHS cycle 1.1 (2000/01) content was adopted from the National Population Health Survey. The list of optional content modules more than doubled for CCHS cycle 2.1 (2003) in response to increased information requirements from stakeholders.

Since 2000, demand for health information has increased significantly, due in large part to requirements for provincially-comparable indicators, resulting from federal/provincial accords on health care funding. In response, sub-sample content was introduced in CCHS cycle 2.1. Common content (asked of all respondents) was reduced by five minutes and replaced by three five-minute content sub-samples of about 30,000 respondents each. This allowed the production of national and provincial data in new high-

priority areas while maintaining the average length of the interview at 45 minutes.

2.2. Sample Design

Since 2000, the CCHS sample design has had to be adapted because of variations in stakeholder requirements. The types of changes range from sample increases to allow the production of estimates for sub-regions, to adapting the sample design to accommodate health region boundary changes. A small area estimation approach has also been developed to help smaller health regions better study less prevalent health conditions. Some of the additional requirements have also involved targeting specific sub-populations, such as nurses and war veterans.

The initial sample design took sample almost equally from an area frame and from a Random Digit Dialling (RDD) frame. Because of low hit rates and other collection difficulties associated with RDD, the sample design was adapted part way through the CCHS cycle 1.1 to select most of the telephone sample from a list frame of telephone numbers. This approach, which yielded much better hit rates, was implemented in all health regions with good coverage from the area frame, which was the case for the majority of regions.

2.3. Responding to Emerging Needs

Because the survey was designed to be very flexible, it has been possible to adapt the survey program based on emerging needs. In cycle 1.1, content was added to the questionnaire in the last quarter of data collection to collect information on patient satisfaction and flu shots in response to requirements by provincial ministries of health. In parallel to this, a supplementary survey on access to health care services, on wait time for specialist visits and some medical services was conducted from a sub-sample of the CCHS. This survey was repeated during cycle 2.1 with a similar design.

2.4. Add-ons for Stakeholders

A number of sample additions were made to the survey to accommodate specific requirements. In cycle 1.1, sample was added by the province of Prince-Edward-Island to allow the production of estimates at the sub-health region level. The same was done for some health regions in Quebec in cycles 2.1 and 3.1. In cycle 2.1, the James-Bay-Cree health region was added to the Quebec sample as a sample buy-in by the region. This region would normally have been out-of-scope in the CCHS design, as it is comprised of Indian

communities (reserves). (Indian reserves are not in scope for the CCHS.)

3. Increasing the CCHS Program's Flexibility

Although several adjustments were made to its initial design, the CCHS is still under increased pressure from its numerous stakeholders to be the main data provider to fill their various data gaps. The demands range from the need for various subjects to be covered in the survey, to the amount of detail that should be collected related to each subject, to the timeliness of the data produced. While the demands on the program exceed its current capacity, it is felt that the current design of the survey is not as efficient as it could be. In order to improve on the efficiency and flexibility of the CCHS, it is proposed to redesign it. Since early 2004, a number of consultation activities have been pursued, including:

- extensive discussion and deliberations within the project team;
- meeting of a special working group of key stakeholders to examine information requirements and design options;
- discussions with federal/provincial/territorial population health surveys advisory committee members;
- meetings with Health Canada and the Public Health Agency of Canada; and
- obtaining advices from the CCHS Steering Committee and Statistics Canada senior management.

In the course of these consultations and deliberations, three aspects of the survey emerged as priorities for a potential redesign:

- the development of an option for continuous collection of the x.1 cycles;
- review of the questionnaire, including how often modules are included and potential for other efficiencies; and
- a review of the data dissemination strategy.

4. Plan for a More Responsive CCHS Program

4.1. Data Collection Approach

In light of the above priorities, it is proposed to conduct the regional component (cycle x.1) of the CCHS under an on-going collection approach using a "rolling sample" design with non-overlapping periodic panels, which can be accumulated over different lengths of time for domains of different sizes. It is

proposed to collect data for each survey cycle (sample of more than 125,000 respondents) continuously over a two-year period. The main elements of the sample design (e.g. sample allocation to health regions, frames, sample selection, weighting and estimation) would be borrowed from the previous x.1 cycles of the CCHS.

Continuous collection would offer two main advantages. First, the ability of the survey to respond to emerging data requirements in a timely manner would be significantly improved when compared with the current design, in which the health region component is in the field only every second year. The main ongoing survey would serve as the vehicle for supplementary surveys which would be conducted either at the time of the main interview or shortly afterwards. Additional filter questions (e.g. households with food insecurity) would be added to the main survey questionnaire to pinpoint specific population profiles for any given special project.

Second, continuous collection would allow Statistics Canada to stabilise the workload of interviewers over time and reduce the peaks and valleys in workload currently experienced. This, in turn, would encourage retention of interviewers and longer-term development of interviewing skills.

The x.2 cycles would continue largely unchanged, although for some x.2 topics (e.g. healthy ageing) there might be potential to benefit from the x.1 cycles as a possible survey frame.

4.2. Survey Content

The primary objective of the redesign of the questionnaire would be to include every year only those core data elements that are either required on an ongoing basis (such as smoking and physical activity) or needed as a correlate in any health analysis (such as disability and socio-demographic characteristics). Other data elements as described below would be collected as optional or theme content. Additionally, more efficient use would be made of sample and interviewing time by including sets of theme content in the questionnaire for 6, 12 or 24 months, depending on analytical and geographic objectives.

The CCHS questionnaire would be reduced in length to 40 minutes and would include the following components:

- optional content, consisting of ten minutes of content selected by each health region from a predetermined list according to local needs

and priorities (this is unchanged from the previous approach);

- core content, consisting of approximately 20 minutes of content asked in all health regions in all cycles of the survey;
- theme content, consisting of related sets of questionnaire topics totalling approximately ten minutes of interview time. In order to manage expectations of data users, a multi-year plan for the inclusion of themes would be developed in consultation with principal data user groups.

To further increase the flexibility of the survey in responding to new requirements, approximately two minutes of interviewing time would be held in reserve to accommodate emerging data needs, which are not anticipated in the multi-year plan for theme content.

4.3. Dissemination Approach

Tabular data and microdata files would be made available within six months of the 12-month and 24-month points of each two-year cycle. The 12-month release would consist of any theme content collected over the first 12-months of the cycle plus core content collected during the same time period. At the 24-month point, two sets of products (tabular and micro data) would be released, the first consisting of theme content collected during months 13-24 plus core content; the other would consist of any theme content collected over months 1-24, core content and optional content collected over months 1-24.

Public-use microdata files (PUMF's) would be produced only at the end of each 2-year cycle. The need to analyse all data collected over that time to mitigate the risk of disclosing the identity of respondents would make it impractical to produce PUMF's after the 12-month point.

In the longer term, possibilities for introducing data products based on accumulation of sample for more than 2 years (eg. core content accumulated over 6 years) will be considered.

5. Sample Design for the CCHS

The primary objective of the CCHS cycle 4.1, which will start collection in either July 2006 or January 2007, will be to provide reliable cross-sectional estimates of health determinants, health status and utilization of the health system for more than 120 health regions. The provincial component under the redesigned program (CCHS cycle 4.2) is scheduled to begin collection in January 2008. The plan potentially

calls for an in-depth survey on healthy ageing of Canadians aged 55 and older, with the objective of producing cross-sectional estimates at the provincial level.

What follows is a detailed description of the proposed sample design for the regional component (cycle 4.1) as well as an overview of the provincial component (cycle 4.2). It is of importance to mention that possible revisions to the proposed design could still be implemented prior to the start of the data collection.

5.1. Sample Design of the Regional Component

5.1.1 Target Population

The CCHS cycle 4.1 will target persons aged 12 years or older living in private dwellings in the ten provinces and the three territories. Persons living on Indian Reserves or Crown Lands, clientele of institutions, full-time members of the Canadian Forces and residents of certain remote regions are excluded from this survey. The CCHS covers approximately 98% of the Canadian population age 12 or older.

5.1.2 Health Regions

For administrative purposes, each province is divided into health regions (HR) and each territory is designated as a single HR (Table 1). Statistics Canada, in consultation with the provinces, has made minor changes to the boundaries of some of the HR's to correspond to the geography of the 2001 Census. The CCHS cycle 4.1 will collect data in 119 HR's in the ten provinces, in addition to one HR per territory, totalling 122 HR's.

5.1.3 Sample Size and Allocation

To provide reliable estimates to the 122 HR's, and given the budget allocated to the CCHS regional component, a sample of 128,200 is targeted. Although producing reliable estimates at the HR level is a primary objective, the quality of the estimates at the provincial level for most key characteristics is also an objective. The proposed allocation strategy described below is basically the strategy implemented to allocate the sample of the CCHS cycle 1.1 for which various scenarios were considered for distributing the sample between the provinces and the HR's. Because the size and number of HR's varied considerably from one province to the other, it was difficult to establish equilibrium between regional and provincial needs. An approach focused primarily on the HR's would have penalized heavily-populated provinces that have a relatively small number of HR's, whereas an

approach focused on the provinces would have produced the opposite effect.

The strategy that was adopted consists of three steps and gives relatively equal importance to the HR's and to the provinces (Béland *et al.*, 2000). The first two steps allocate the sample between the provinces as a function of their respective populations and of the number of HR's they contain. The three territories are not included in these first two steps and are treated separately. Table 1 provides a detailed distribution of the regional component sample size by province.

Table 1 – Provincial Sample Sizes

Province	Number of HR's	Total Sample
Newfoundland/Labrador	4	4,010
Prince Edward Island	4	2,000
Nova Scotia	6	5,040
New Brunswick	7	5,150
Quebec	16	24,280
Ontario	36	41,760
Manitoba	10	7,500
Saskatchewan	11	7,720
Alberta	9	12,200
British Columbia	16	16,090
Yukon	1	850
Northwest Territories	1	900
Nunavut	1	700
Canada	122	128,200

The third and last step consists of allocating each provincial sample among the provincial HR's proportionally to the root of the size of the population in the regions. This three-step approach guarantees each HR sufficient sample. Table 2 summarizes the distribution of HR's by population size, as well as the average sample size planned by HR category.

Table 2 – Mean Sample Sizes by Category of HR's

Category of HR's	Population Size	# of HR's	Mean Sample Size
Small	Less than 75,000	37	600
Medium	75,000 - 240,000	49	900
Large	240,000-640,000	28	1,500
Very Large	640,000 and +	8	2,600

These sample sizes are inflated before data collection to take into account out-of-scope dwellings, vacant

dwellings and anticipated household level non-response.

5.1.4 Strategies for Sampling Households

The choice of a sampling frame depends on many factors but, first and foremost, the frame must correspond as much as possible to the target population of the survey. Moreover, the creation, utilization, updating and verification of the sampling frame must fall within the operational and budget constraints of the survey. As for the first three regional components of the CCHS (cycles 1.1, 2.1 and 3.1), the CCHS cycle 4.1 will make use of three sampling frames to select the sample of households.

For 117 health regions in the ten provinces, two overlapping sampling frames, the area frame established for the Canadian Labour Force Survey (LFS) and a list frame of telephone numbers, will be used to select the sample. For each health region, half the sample of households will be selected from the area frame, while the other half will come from the list frame of telephone numbers. In two remote health regions (northern Quebec and northern Saskatchewan), an RDD frame of telephone numbers will be used to select the sample (the utilization of the area frame in those two regions would generate extreme collection costs). In the three territories, an area frame will primarily be used to select the sample and an RDD frame will be used in Whitehorse (Yukon) and Yellowknife (Northwest Territories) to complement the sample.

Apart from the fact that the target population is the same as that of the LFS, the advantages of using the area frame set-up for the LFS for selecting the sample are undeniable. The infrastructure, which is already in place for updating new buildings, demolished buildings and excluded units, as well as the entire evaluation process of the frame coverage, are definite assets. Moreover, given that several other Statistics Canada household surveys also use this area frame, it is easier to control sample overlap between surveys. For various reasons, but primarily in consideration of response burden, it was decided to choose a sample of new dwellings and not a sample of rotated-out dwellings from the LFS. (It is worth noting that, for the LFS, the dwellings selected remain in the sample for six consecutive months and are then replaced. These rotated-out dwellings are then available for other surveys.)

The limitations to using telephone frames (list and RDD) are evident: under-representation of households without telephones or with cellular phones only,

unlisted phone numbers (for the list frame), the generally lower response rate and the need to make several calls before contacting a valid household. However, in order to balance costs and operational constraints, the sample of households for the CCHS cycle 4.1 will be selected from the above three sampling frames. Next we describe in more details the sampling strategy for each frame.

5.1.5 Sampling from the Area Frame

The CCHS will use the area frame designed for the LFS as a sampling frame to select half the sample of households for most health regions in the ten provinces. The sampling plan of the LFS is a multistage stratified cluster design in which the dwelling is the final sampling unit (Statistics Canada, 1998). In the first stage, homogeneous strata are formed and independent samples of clusters are drawn from each stratum. In the second stage, dwelling lists are prepared for each cluster and dwellings, or households, are selected.

For the purpose of the LFS sampling plan, each province is divided into three types of regions: major urban centres, cities and rural regions. Geographic and/or socio-economic strata are created within each major urban centre. The other cities and rural regions of each province are stratified first on a geographical basis, then according to socio-economic characteristics. Some urban centres have separate strata for apartments or for census dissemination areas in which the average household income is high. In each stratum, six Primary Sampling Units (PSU's) or clusters are chosen by a random sampling method with a probability proportional to size (PPS), the size of which corresponds to the number of households. Once the new clusters are listed, the final LFS sample is obtained using a systematic sampling of dwellings.

Requirements specific to CCHS mean that some modifications must be made to this sampling strategy. To get a base sample of 62,000 responding households for CCHS (ten provinces), 84,000 dwellings must be selected from the area frame (to account for vacant or out-of-scope dwellings and non-responding households). On an on-going monthly basis the LFS design provides approximately 68,000 dwellings distributed across the various economic regions in Canada whereas the CCHS required a total of 84,000 dwellings distributed in the HR's, which have different geographic boundaries from those of the LFS economic regions. Overall, the CCHS required 24% more dwellings than those generated by the LFS selection mechanism, or an *adjustment factor* of 1.24 (84,000/68,000). At the HR level, however, the

adjustment factors vary from 0.6 to 6.0, which require further adjustments.

The changes made to the selection mechanism in a HR vary depending on the size of the adjustment factors. For HR's that have an adjustment factor less than or equal to 1, a sub-sample is selected. For those with a factor greater than 1, but less than or equal to 2, the sampling process of dwellings within a PSU is repeated for all selected PSU's that are part of the same HR. For HR's with a factor greater than 2 but less than or equal to 4, the PSU sampling process, as well as that of dwellings in a PSU, is repeated. For HR's with a factor between 4 and 6, the PSU sampling process is repeated not once but twice, while that of dwellings is repeated only once. Where the chosen approach creates an unnecessary surplus of dwellings, a sub-sample is selected. It should be noted that the changes made to the LFS mechanism result in, at most, tripling the number of PSUs selected and, at most, doubling the number of dwellings selected in the PSUs, which explains the maximum adjustment factor of 6.0.

5.2 Sampling of Households in the Three Territories

For operational reasons, the area frame sample design implemented in the three northern territories will have one additional stage of selection. For each territory, in-scope communities are first stratified based on various characteristics (population, geography, percent Inuit and/or Aboriginal and median household income). There are five design strata in Yukon, ten in the Northwest Territories and six in Nunavut. Then the first stage of selection consists of randomly selecting one community with a probability proportional to population size within each design stratum. From that point on, the household sampling strategy from the area frame, within the selected community is identical as the one described above – the secondary sampling units are the clusters and the tertiary sampling units are the dwellings. It is worth mentioning that the frame for CCHS covers 90% of the private households in Yukon, 97% in the Northwest Territories and 71% in Nunavut.

5.2.1 Sampling from the Telephone List Frame

The list frame of telephone numbers will be used in all but five HR's (the two RDD-only HR's and the three territories) to complement the area frame. The InfoDirect database, a commercially available product consisting of names, addresses and telephone numbers from telephone directories in Canada, is linked to internal administrative conversion files to obtain postal

codes and these are mapped to HR's to create list frame strata. There is one list frame stratum per HR. Within each stratum the required number of telephone numbers is selected using a simple random sampling process from the list. Additional telephone numbers are selected to account for the numbers not in service or out-of-scope. The hit rates observed under the list frame approach vary from 80% to 90%.

It is important to mention that the coverage of the list frame is less than the one for RDD as unlisted numbers do not have a chance of being selected. Nevertheless, as the list frame will only be used in HR where the area frame is also used, the impact of the under-coverage of the list frame will be minimal and dealt with at the weighting stage.

5.2.2 Sampling from the Telephone RDD Frame

In two HR's (northern Quebec and northern Saskatchewan), and in Whitehorse and Yellowknife, a RDD sampling frame of telephone numbers will be used to select the sample of households. The sampling of households from the RDD frame uses the Elimination of Non-Working Banks (ENWB) method, a procedure adopted by the General Social Survey (Norris and Paton, 1991). A hundreds bank (the first eight digits of a ten-digit telephone number) is considered to be non-working if it does not contain any residential telephone numbers. The frame begins as a list of all possible hundreds banks and, as non-working banks are identified, they are eliminated from the frame. It should be noted that these banks are eliminated only when there is evidence from various sources that they are non-working. When there is no information about a bank, it is left on the frame. The InfoDirect database and telephone companies' billing address files will be used in conjunction with various administrative files to eliminate non-working banks.

Using available geographic information (postal codes), the banks on the frame are regrouped to create RDD strata to encompass, as closely as possible, the HR areas. Within each RDD stratum, a bank is randomly chosen and a number between 00 and 99 is generated at random to create a complete, ten-digit telephone number. This procedure is repeated until the required number of telephone numbers within the RDD stratum is reached. The hit rates observed in those areas are between 15% and 20%.

5.3 Sampling of Persons

Interviewing more than one person in the same household allows for a reduction in the cost of collection, since a large part of these costs are

attributable to the process required to reach the household. However, strong similarities observed among members of the same household can lead to an undesired cluster effect for certain important survey characteristics. Moreover, and probably the most important disadvantage considering the length of the interview, the response burden of the household is increased.

On the other hand, selecting only one person per household represents an increase in collection costs, since a greater number of households must be sampled. Also, as the chances of being part of a sample are inversely proportional to the number of persons in the household, certain age groups are either under- or over-represented. In particular, selecting only one person per household, under-represents persons coming from large households (typically parents and children), and over-represents persons coming from small households (often single people and the elderly).

In CCHS cycle 1.1, one person per household was selected for 86% of the sampled households and two persons were chosen in the remaining ones. For various reasons but mainly for operational purposes, that strategy was not repeated in cycles 2.1 and 3.1, in which only one person per household was selected using age-based varying probabilities of selection. (Person-level non-response was twice as high in two-person households as in one-person household in cycle 1.1.) Although the final parameters have not yet been determined, it is proposed, for the CCHS cycle 4.1, to adapt and implement the person-level sampling strategy designed for the cycle 3.1, where one person per household was selected using varying probabilities.

Table 3 compares the age group distributions of the targeted population from the 2005 Census projections and a simulated CCHS sample, where one person per household is selected with equal probability. The results clearly demonstrate that, by selecting one person per household with equal probability, the 12-to-19 age group is greatly under-represented while old persons are over-represented.

Table 3 – Percent Distribution by Age Group

Age Group	2005 Census Projections	CCHS sample (equal prob.)
12-19	12.4	8.2
20-29	16.0	12.3
30-44	26.5	24.7
45-64	30.6	33.4

65+	14.5	21.5
-----	------	------

The selection of individual respondents for the CCHS cycle 4.1 is designed to consider user needs, cost, design efficiency, response burden and operational constraints. Several scenarios, using various parameters, were simulated with the objective of identifying an optimal approach that would guarantee sufficient number of individuals in each age group, without generating extreme sampling weights.

Table 4 gives the selection weight multiplicative factors used to determine the probabilities of selection of individuals in sampled households by age. As an example, for a three-person household (two age 45-64 adults and one 15-year-old), the teenager would have 5 more chances of being selected than the adults. In order to keep the number of extreme weights to a minimum there is one exception to this rule: if the size of the household is greater or equal to 5 or the number of 12-19 is greater or equal to 3 then the selection weight multiplicative factor equals 1 for all individuals in the household.

Table 4. Selection Weight Multiplicative Factor by Age

Age	Selection Weight Multiplicative Factor				
	12-19	20-29	30-44	45-64	65+
Factor	5	2	2	1	1

Table 5 shows the expected distribution of a simulated CCHS sample when selecting one person using the rule described in Table 4 in all sampled households. The results show that the 12-to-19 age group representativeness is improved without penalizing greatly the other age groups.

Table 5 – Expected CCHS Sample Distribution by Age Group

Age Group	2005 Census projections	CCHS sample (unequal prob.)
12-19	12.4	12.6
20-29	16.0	12.9
30-44	26.5	23.9
45-64	30.6	29.5
65+	14.5	21.1

5.3.1 Data Collection Specifics

As discussed earlier, the data collection of the CCHS cycle 4.1 is scheduled to begin in either July 2006 or January 2007 and will span 24 months. This should allow for both an even distribution of interviewer

workload and the elimination of possible seasonal effects on a yearly basis. In each health region (HR), the entire sample will be divided in 12 non-overlapping 2-month collection periods where each 6-month sample (three collection periods) will be representative nationally and provincially. This should allow for great flexibility, over the course of a given cycle, to adjust sampling factor assumptions (such as response rates and telephone hit rates), to correct computer-assisted interviewing application errors or to introduce new or buy-in content. Two weeks prior to a collection period, introductory letters, describing the importance of participating in the survey, will be sent to all dwellings (area and telephone frames), for which a valid mailing address is available. All area frame cases will be sent to field interviewers and the telephone frame cases will be assigned to interviewers working in call centers.

As was done for previous x.1 cycles of the CCHS, the field interviewers will be instructed to find the dwelling addresses, assess the status of the dwellings (out-of- or in-scope) and list all household members to allow for the random selection of one individual aged 12 or older. If the selected individual is available at that time then the interviewer will conduct a personal interview. If not then the interviewer will have the option of coming back at a later date for a personal interview or completing the interview over the phone. In previous regional components, between 30% and 40% of the area frame cases were completed over the phone and it varied from 0% to 70% among HR's (only a few HR's had more than 50% interviews completed over the phone).

For the telephone frame cases, the call center interviewers will be instructed to assess the status of the phone numbers, list all household members and conduct an interview with the selected individual at that moment or at a later date.

It is important to mention that, in order to preserve comparability over time at the health-region level, an effort is made to maintain the same mix of area frame/telephone frames and personal/telephone interviews from one cycle to another. A study on collection mode effects (personal versus telephone) conducted as part of CCHS cycle 2.1 (St-Pierre and Béland, 2004) showed significant differences between estimates for key health characteristics.

5.3.2 Weighting and Estimation

Taking into account the sample design, each respondent will be assigned a sampling weight to represent his or her contribution to the total

population. The sampling weights will be used to derive estimates for all characteristics surveyed during the time period. Previous experiences with the first three cycles of the CCHS will serve as a basis to develop and implement the weighting and estimation strategy for the redesigned CCHS (Brisebois and Thivierge, 2001).

The combination of the health region-level sample design and the data collection strategy will allow for the production of sampling weights and estimates at various points in the two-year cycle. It is important to mention that separate sampling weights will be derived each time microdata files are produced.

Because the CCHS uses two overlapping sampling frames with separate sample designs for most parts of the country, two weighting strategies with various adjustments will be processed side-by-side and then integrated using a dual-frame technique. The integrated weights will then be adjusted to control for seasonal effects and calibrated to population projections using a one-dimensional post-stratification adjustment of 10 age/sex post-strata (that is, the age groups 12 to 19, 20 to 29, 30 to 44, 45 to 64 and 65 or older for each sex) within each health region. Before the integration adjustment, key factors will determine the weighting strategy for the CCHS data. For the area frame units, these factors include:

- use of stratified, multistage design, involving probability sampling proportional to size at all stages except the final stage, when systematic sampling of dwellings are used;
- bi-monthly stabilization of sampled dwellings;
- household-level non-response;
- selection of one person per household using age-based varying probabilities of selection;
- person-level non-response.

For the list frame sampling units, some of the determining factors include:

- use of SRS of telephone numbers within each list frame stratum;
- household-level non-response;
- households not included in the frame because of no telephone line or not listed;
- selection of one person per household using age-based varying probabilities of selection;
- person-level non-response.

As was done for the previous cycles of the CCHS, it is proposed that variances will be obtained using the

bootstrap re-sampling approach, similar to the bootstrap adaptation used for the National Population Health Survey (Yeo, Mantel and Liu, 1999).

5.4. Sample Design of the Provincial Component

The primary objective of the CCHS cycle 4.2 (2008) will be to produce cross-sectional estimates on factors related to health and ageing of the Canadian population at both provincial and national levels from a sample of approximately 30,000 respondents. It is proposed that the target population be Canadians aged 55 and over, and interviews will be conducted in person. The possibility of including the population residing in health institutions will also be considered. The various elements of this survey will be determined in collaboration with key stakeholders in the coming years. Data collection is expected to begin in January 2008 and will extend over 12 months. Both questionnaire content and sample design for this provincial component of the CCHS are still under development at the time of writing this paper.

6. Future Directions

Currently, population health and health care services are some of the most important concerns in Canadians' minds and are expected to remain a top priority for decision makers and health care researchers for years to come. The addition of the CCHS to the existing National Population Health Survey in 2000 made the Statistics Canada health survey program more comprehensive and robust.

The proposed changes to the CCHS should help reduce the number of information data gaps in the short- or the mid-term. However, in the long term, the CCHS program will have to remain flexible over time not only to adapt to new or changing requirements, but also to react to findings from the other health survey programs at Statistics Canada, and help paint the most complete picture of the state of health of the Canadian population.

Using the data from this health survey program, policy makers and health care professionals will be able to identify benchmarks and track progress. Therefore, this should help them take appropriate action towards addressing the shortcomings of the health care system in Canada.

Acknowledgements

The authors would like to thank all project team members who have worked on various aspects of the

CCHS redesign. A special thank to Larry MacNabb and Steven Thomas for their helpful comments during the writing of this paper.

References

- Béland, Y., Bailie, L., Catlin, G. and Singh, M.P. (2000). CCHS and NPHS — An Improved Health Survey Program at Statistics Canada, *2000 Proceedings of the American Statistical Association Meeting, Survey Research Methods Section*, Indianapolis.
- Brisebois, F. and Thivierge, S. (2001). The Weighting Strategy of the Canadian Community Health Survey, *2001 Proceedings of the American Statistical Association Meeting, Survey Research Methods Section*, Atlanta.
- Canadian Institute for Health Information (1999a), Health Information Roadmap: Beginning the Journey. (1-895581-32-X).
- Canadian Institute for Health Information (1999b), Health Information Roadmap: Responding to Needs. (1-895581-30-3).
- Norris, D.A. and Paton, D.G. (1991). Canada's General Social Survey: Five Years of Experience, *Survey Methodology*, 17, 227-240.
- Statistics Canada (1998). *Methodology of the Canadian Labour Force Survey*. Statistics Canada. Cat. No. 71-526-XPB.
- St-Pierre, M. and Béland, Y. (2004). Mode Effects in the Canadian Community Health Survey: A Comparison of CAPI and CATI, *2004 Proceedings of the American Statistical Association Meeting, Survey Research Methods Section*, Toronto.
- Yeo, D., Mantel, H. and Liu, T.P. (1999). "Bootstrap Variance Estimation for the National Population Health Survey", *2005 Proceedings of the American Statistical Association Meeting, Survey Research Methods Section*, Baltimore.