Challenges in the Design of the Canadian Community Health Survey on Healthy Aging

Leon Jang[1], Martin Provost[1], Adam Sherk[1]
[1]Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, Ontario, K1A 0T6, Canada

Abstract
The Canadian Community Health Survey on Healthy Aging (CCHS–HA) was conducted by Statistics Canada between December 2008 and November 2009. The survey focused on the health of Canadians aged 45 and older by examining the various factors that impact healthy aging. A nationally representative sample of 30,865 respondents completed personal interviews which included an interactive component on cognition measures. The sample was selected using a multi-stage design that used the 2006 Census of Population as a frame in order to target the population of older Canadians. A number of sampling and estimation issues are presented in this paper including the use of the census as a frame, clustering, hard-to-reach populations and cognitive measures during interviewing. The effectiveness of the design is also examined using the results from the survey.

Key Words: Census frame, clustering, health survey, cognitive interviewing

1. Introduction

The age of the average person living in Canada is rapidly increasing. In 2009, the median age of Canada's population was 39.5 years, an increase of 3.1 years from 1999. Some of the reasons for this include the increasing longevity of Canadians, decreasing fertility rates and the advancing age of people from the baby boomer generation born between 1946 and 1965. According to projections by Statistics Canada (2010a), the number of senior citizens aged 65 years and older is expected to double over the next 25 years, from 4.7 million seniors in 2009 to between 9.9 and 10.9 million seniors in 2036.

With most of the baby boomers currently in the 45 to 64 age group, this generation accounts for 40% of Canada's working-age population aged 15 to 64 years, which is the highest proportion to date. This demographic will soon transition into retirement and, if current trends continue, this transition will come at an earlier age than in the past. According to Statistics Canada's Labour Force Survey, the average age of retirement in 2009 was 62, a full three years younger than the average age 20 years ago. Health-related habits may have an effect on retirement decisions, so public health programs could potentially play a role in the participation of older workers in the labour force (Park 2010).

Health policy makers and researchers understand the importance of planning for the future of an aging population in the coming years. The demographic shift will necessitate changes in decisions relating to health care, social services and income support programs. For example, it will be important to promote lifestyle changes such as good nutrition and

regular physical activity in order to reduce the risk of health problems that could eventually impact the health care system.

Statistics Canada's Canadian Community Health Survey on Healthy Aging was designed to collect information about the factors, influences and processes relating to healthy aging. The survey focuses on Canadians aged 45 years and older and looks at several factors including general health and well-being, physical activity, the use of health care services, social participation and transitions from work to retirement. In creating the survey design, there were various challenges that were unique to the specific population of interest. Even with sound planning, issues were encountered during and following data collection that had to be resolved to ensure that results were relevant and accurate.

In section 2, a brief overview of the Healthy Aging survey is given. Section 3 describes the challenges involved with the selection of the 2006 Census of Population as a frame, including both the benefits and drawbacks of its use. The reasons for cluster sampling are presented in section 4, along with the methods used to construct clusters and the criteria used to optimize their construction. In section 5, the issues involving two hard-to-reach populations, respondents living in rural areas and Canadians aged 75 and older, are outlined as well as the sample design used to resolve them. Section 6 discusses the challenges encountered during collection and the measures taken to improve survey response. Other issues involving the cognition and physical activities modules are described in section 7. In section 8, an evaluation of results from the survey is presented and, finally, some concluding remarks are provided in section 9.

2. The Canadian Community Health Survey on Healthy Aging

Statistics Canada's CCHS–HA is the latest focused content survey conducted as part of the Canadian Community Health Survey (CCHS) program. The CCHS focused content component is designed as a cross-sectional survey to provide results for the 10 Canadian provinces on specific health topics. Past survey topics were mental health in 2002 and nutrition in 2004.

Planning for the CCHS–HA began in January 2006. Primary stakeholders were surveyed to determine their upcoming data needs. The consultation process involved stakeholders from Health Canada, the Public Health Agency of Canada, representatives from the provinces and their health regions as well as members of the research community. Participating stakeholders were asked to prioritize various themes in terms of relevance, issues, timing and topics within themes. As a result of this consultation, healthy aging was chosen as the focused content of the survey to be conducted between December 2008 and November 2009.

The CCHS–HA has four specific survey objectives (Statistics Canada 2010b):

- To better understand the aging process of Canadians aged 45 and over by collecting data on various aspects of their health and well being, use of health care services, social support and participation and transitions from work to retirement;
- To examine how lifestyle determinants affect health as people age;
- To examine the links between healthy aging and social, demographic, geographic and economic variables or characteristics using a multidisciplinary approach; and

- To provide information on successful aging by age group and sex.

The survey questionnaire was conducted using computer assisted personal interviews. Excluding the entry and exit modules, it consisted of 37 modules which are listed below. A summary description of these modules can be found in the survey user guide (Statistics Canada 2010b).

- Proxy interview
- General health
- Sleep
- Height and weight – self-reported
- Chronic conditions
- Health utility index
- Pain and discomfort
- Satisfaction with life scale
- Cognition
- Physical activities
- Nutritional risk
- Oral health
- Medication use
- Dietary supplement use – vitamins and minerals
- Smoking
- Alcohol use
- Changes made to improve health
- Falls
- Instrumental activities of daily living
- Basic activities of daily living
- Health care utilization
- Care receiving 1
- Care receiving 2
- Social support availability
- Social participation
- Care giving
- Care giving expenses
- Depression
- Loneliness
- Transportation
- Labour force
- Reasons for retirement
- Retirement planning
- Home owner
- Income
- Socio-demographic characteristics
- Administration information

The target population of the CCHS–HA was persons aged 45 years and older who are living in private dwellings in the 10 provinces. Residents of the three Canadian territories, persons living on Indian reserves or Crown lands, those residing in institutions, full-time members of the Canadian Forces and residents of certain remote regions were excluded from the survey.

To meet the survey objectives and given the budget allocated to the survey, a sample size of 32,000 respondents was desired over a period of one year. The goal was to produce reliable estimates by province, age and sex using five age groups (45–54, 55–64, 65–74, 75–84 and 85 and older).

3. Use of the census as a frame

Several sampling frames were considered. In terms of sampling, the main challenge was to obtain enough respondents in the older age groups. Targeting rare populations, such as individuals aged 85 or older, by sex in a given province creates a challenge because of the scarcity of available frames providing adequate coverage of the population of interest.

3.1 Non-census sampling frames considered

Several sampling frames aside from the census were considered, but not one of these, on its own, met the requirements of the Healthy Aging survey. The options considered were as follows:

- Area frame of the Labour Force Survey (LFS): This frame doesn't provide information on household composition. A sample of over 200,000 households would have been needed to reach the required representation of the 85+ population, which was not a feasible option due to cost constraints.
- LFS outgoing rotation groups: The household composition is available. However, owing to prior commitments, the LFS could only guarantee access to ten outgoing groups at the time of frame evaluations. Desired sample sizes for the 75–84 and 85+ cohorts would not have been achieved.
- Sample from the regional component of the CCHS: The household composition is known, but this option would not have yielded a sufficient number of respondents in any cohort.
- Administrative databases: Several administrative databases, such as the Old Age Security Pension database and the Personal Master file (an income tax list frame), were considered, but none of them constituted a viable option.
- Combining respondents from the regional component of the CCHS with LFS outgoing rotation groups: This also would not have yielded enough respondents.

3.2 The proposed use of the census

In light of the fact that no other frame, or combination of frames, was in a position to provide enough respondents to satisfy requirements by province, age and sex at a reasonable cost, a proposal was put forward to use the census as a sampling frame. Several options for such a use were examined, and using the 2006 Census as the sole frame for the selection of dwellings was identified as the best course of action. The only census information required for these purposes was the date of birth and sex of every resident of each dwelling at the time of the census, as well as the address and other geographical information to be used for clustering.

This information made stratification and clustering of the households possible, as well as identifying households with a high probability of containing at least one member of the targeted age groups. As there were reasons to expect that the composition of many households would have changed between 2006 and 2008, a pilot survey was used to assess changes to household composition. This was required to determine a sample allocation that would yield the desired representation from each age group despite the absence of up-to-date information on household composition at the time of selection. The pilot survey was conducted in November 2007 using respondent households from the area frame of the 2005 CCHS Cycle 3.1, since permission had not yet been obtained to use the census. Information on household composition was two or three years old, mirroring the expected age of the census data at the time the main survey was to be completed.

Unlike the other frames considered, using the 2006 Census provided a sufficient number of respondents, even for the oldest age groups. It also ensured an adequate coverage of the target population. Furthermore, the use of a household frame meant that no tracing was required. The telephone numbers provided in the census provided an additional method to contact the dwellings other than personal visits. Finally, using the 2006 Census

instead of the LFS or CCHS provided greater flexibility for other surveys to use the LFS or the CCHS as their frames.

The main disadvantage of this approach was its incomplete (though adequate) coverage, owing partly to the mobility of households and individuals. Only addresses containing an individual aged 43 years or older at the time of the 2006 Census were eligible for selection as this would make them 45 years or older at the time of the Healthy Aging survey. This meant that dwellings that were missed, newly built, vacant or occupied only by individuals younger than 43 years at the time of the census might now house members of the survey's target population and yet not be covered by it. Consideration was given to adding a sample from these dwellings but, given the high proportion of them that would have been classified as out of scope and the costs associated with in-person collection, this course of action was not deemed practical.

3.3 Other potential uses of the census

Other potential uses of the census as a sampling frame were considered. One consisted of using it to directly sample individuals. This would have resolved most of the coverage problems described in the previous section (except for new arrivals). Mobility would no longer affect coverage directly, but it might create challenges at the level of tracing, and thus non-response, which could arise for instance in the case of missing or fictitious names. The tracing stage would have been necessary and costly. For budgetary reasons, using the census as a household frame was deemed preferable.

Rather than use the census as the sole household frame, another possibility was to combine it with LFS outgoing rotation groups. For every age group, one part of the sample would come from each frame, with the proportion from the census increasing with age. After balancing the pros and cons of the LFS outgoing rotation groups method against using the census as a sole frame, it was decided that the benefits realized from monopolising ten LFS outgoing rotation groups would be marginal in light of the inconvenience to other surveys.

4. Clustering

With census households settled as the frame, a three-stage sample design was developed for this survey. In the first stage, geographical regions, or clusters, were selected. In the second stage, a sample of households was selected within each cluster. Finally, the last stage consisted of choosing a respondent within each household. Each of these stages is explained in greater detail in the following sections.

Even though the 2006 Census provided a list frame for sample selection, it was necessary to sample at the cluster level in order to control the cost of data collection by means of personal interviews. The challenge was to create clusters containing a sufficient number of potential respondents in the oldest age groups while controlling cluster size. Too small to be used directly, the LFS clusters were also not optimal to use as building blocks for constructing larger clusters. Consequently, it was decided to create customized clusters on the basis of 2006 Census blocks. With the collaboration of Statistics Canada's Geography Division, the Generalized Area Delimitation System (GArDS) software was used to create these clusters.

Prior to creating the clusters some blocks were identified as being isolated, i.e., remote and difficult to reach, using block exclusion criteria similar to those applied by the LFS. The urban/rural status of each block, as established by the 2006 Census, was also obtained since the criteria for cluster formation were not the same in urban and rural environments. In this fashion, clusters were created within each province–urban/rural status combination. These combinations constituted the strata for the first sampling stage. GArDS grouped contiguous blocks using the following criteria (on the basis of 2006 Census numbers):

- At least 100 occupied private dwellings in a cluster;
- A minimum number of dwellings with at least one member aged 85 or older, and a similar criterion for the 75–84 cohort. The actual number varied by province and urban/rural status;
- Cluster size not to exceed a certain limit (in km^2)—this also varied by province and urban/rural status; and
- Emphasis on ease of access (ability to move from one point in the cluster to any other without leaving it) and compactness.

A different weight was assigned to each of these four criteria. Not all clusters satisfied all the criteria, but GArDS created a set of clusters that came closest to meeting these requirements. In total, approximately 17,000 clusters were created.

In order to calculate how many clusters to sample it was necessary to determine the number of households to select within each cluster. This involved trading off cost efficiency against precision. Thus, it was established that a raw sample of 35 units per urban cluster and 20 units per rural cluster was reasonable. Cluster selection was performed on the basis of sampling with probability proportional to size (number of dwellings with individuals aged 43 or older at the time of the 2006 Census) according to the Hanurav-Vijayan algorithm (Vijayan 1968). With this method, the more dwellings with individuals aged 43 or older there were in a cluster according to the 2006 Census, the greater its probability of being selected. This method proved necessary, given the variability in cluster size.

5. Hard to reach populations

In planning the sample, there were two specific populations that were anticipated to be more challenging. First were those living in areas defined as being rural according to the census block. It was expected that personal interviews here would be more costly given that travel to these places would be more difficult. Second, the population of Canadians aged 75 years and older was limited, comprising just 10.8% of the 2006 Census population aged 45 and older. For the 85 and older population, the proportion of the target population is 3.6%.

It was important to give special consideration to these two populations to ensure that they were adequately represented in the sample in order to have sufficient numbers for later analyses. Therefore, the sample design was planned around user needs as well as cost, response burden and operational constraints.

5.1 Sample size and allocation

To meet the survey objectives specified in section 2 for each province and given the budget allocated to the survey, a sample of 32,000 responding units was desired over a period of one year. A two-step strategy was used to allocate the sample to the provinces. First, 125 sample units were allocated to each domain of interest (10 age/sex groups) in each province for a total of 12,500. The remaining 19,500 units were allocated to the provinces using a power allocation method with power q=0.7 (Bankier 1988). The total sample size of any given province was found by adding the number of units obtained in the two steps.

In order to have a good urban and rural representation in each province, the sample was subsequently allocated to two strata: urban and rural, as derived from the 2006 Census blocks. The provincial sample was proportionally allocated to the two strata according to the number of dwellings having people aged 45 and older in each stratum. Sample sizes were then inflated before data collection to take into account an anticipated out-of-scope rate of 24% and a non-response rate of 20%. In particular, the out-of-scope adjustment was important in this survey since selected households would be rejected if there was no one within the target population (aged 45+) living there. The raw sample size needed to obtain 32,000 respondents was estimated at 52,010. Table 1 gives the raw sample sizes for the Healthy Aging survey by province and urban/rural status.

Table 1: Raw Sample Sizes by Province and Urban/Rural Status

Province	**Urban**	**Rural**	**Total**
Newfoundland and Labrador	1,885	1,320	3,205
Prince Edward Island	1,397	1,160	2,557
Nova Scotia	2,201	1,560	3,761
New Brunswick	2,016	1,440	3,456
Quebec	7,490	1,316	8,806
Ontario	9,762	1,280	11,042
Manitoba	3,045	740	3,785
Saskatchewan	2,555	1,079	3,634
Alberta	4,585	675	5,260
British Columbia	5,904	600	6,504
Canada	**40,840**	**11,170**	**52,010**

The sample was divided into six collection periods of two months to allow for flexibility with the stratification parameters described in section 5.2 and the sampling parameters described in section 5.3. In addition, the samples for both the December–January and February–March periods were reduced due to decreased interviewer resources at those times. This sample reduction was reallocated to the other four collection periods.

5.2 Selection of households

As described in section 3, the 2006 Census of Population was chosen as the sampling frame because it was able to target dwellings that were more likely to include Canadians in the 75–84 and 85+ age groups. In each selected cluster, the in-scope dwellings were divided into three strata:

- Those with at least one person aged 85 and older;
- Those with only people under age 55 (but at least one aged 45–54); and

- All other dwellings.

The first two strata were created to ensure that the number of people selected in each age group was appropriate. Without this stratification, the sample would have included too many people in the 45–54 age group and not enough in the 85+ group.

The sample of 35 dwellings in urban areas and 20 in rural areas was then allocated to each of these three strata, with an allocation that varied by province. As previously mentioned, one challenge of using this method was that the sample would be selected based on the household composition at the time of the 2006 Census, occurring 2.5 to 3.5 years prior to collection for this survey. The composition of many households would have changed during that time. To determine how the sample should be allocated to the three strata, simulations were run using the household composition at the time of the census and changes in household composition were imputed based on results from the pilot survey of 1,000 respondents conducted in 2007. The number of dwellings required in each stratum was obtained using simple random sampling.

5.3 Selection of interviewees

One person in each dwelling was chosen as the respondent for the survey. Upon visiting a selected dwelling, the household composition at the time of the survey was obtained. Households with no eligible respondents (those with only people aged 44 and younger or those not in the target population) were classified as out of scope. For the other households, one respondent was selected at random among all eligible respondents according to varying selection probabilities.

Every household member aged 45 and older was assigned a selection probability factor according to the five age groups for which estimates were required (45–54, 55–64, 65–74, 75–84, 85+). The selection probabilities varied by province in order to achieve the targeted number of respondents in each age group as outlined above in Table 1. A simulation study using household composition information from the pilot survey was used to create the selection probability factors.

Once the respondent was selected, the interview was required to be completed within the assigned two-month collection period. For example, in the December 2008 to January 2009 collection period, the interview had to be finished by January 31, 2009 no matter when the selection of the respondent occurred.

6. Challenges during collection

Collection began in December 2008, and the response rate was monitored by province and age group. During the first collection period, there were three trends that were observed. First, the out-of-scope rate for the December–January period was 18.4%, which was lower than the expected rate of 24%. Also, the non-response rate for the collection period was 24.0%, which was higher than the expected rate of 20%. Finally, the targeted number of respondents in the first collection period was met in the 45–54 and 55–64 age groups but not in the 65–74, 75–84 and 85+ age groups.

6.1 Corrective measures

In reaction to these trends, two specific corrective measures were implemented. First, the selection probability factors described in section 5.3 were adjusted beginning with the

April–May collection period in order to give a higher chance of selection to the older age groups. The adjustment was tailored to the age group distribution of responses in each province. In Newfoundland and Labrador, the selection probability factors were reduced for the 45–54, 55–64 and 65–74 age groups; in Prince Edward Island, the factors were reduced for only the 45–54 age group; and in the other eight provinces, the factors were reduced for the 45–54 and 55–64 age groups. To limit the possibility of extreme weights, the ratio between the largest and smallest factors in each province was fixed at a maximum value of nine.

Second, the sample allocation among the three strata presented in section 5.2 was revised starting with the June–July collection period. This revision was based on what was observed in the first two collection periods and resulted in more dwellings chosen from the first stratum (those with at least one person aged 85 and older) and fewer chosen from the second stratum (those with only people under age 55). A second revision was done for the October–November collection period for the provincial sample in Nova Scotia to increase the number of dwellings chosen from the second stratum.

6.2 Resent cases

As mentioned in section 5.3, interviews were required to be completed during an assigned collection period. For each two-month collection period, regional collection offices were instructed to use the first four weeks to resolve the majority of the sample, with the next four weeks being used to finalize the remaining sample and to follow up on outstanding non-response cases. In addition, all cases were to have been attempted by the second week.

A review of status codes following the end of each collection period found that a number of non-response cases had been attempted but personal contact was not achieved. This proved to be somewhat problematic given that these cases remained in scope and were only terminated due to the constraint of the two-month collection period. As well, some cases were abandoned and coded as a non-response because the selected respondent normally living at the dwelling was absent for the duration of the collection period.

During the second half of survey collection, the decision was made to resend these types of non-response cases back into the field in order to improve on the number of completed responses. Non-response cases were only to be resent if they had one of three final status codes:

- No one home or no answer;
- Interview prevented due to weather conditions; or
- Absent for duration of survey (collection period).

Since the survey was scheduled to end in November 2009, resends could only occur for cases in the first five collection periods. A total of 1,163 cases from the December–January, February–March and April–May collection periods were resent into the field on August 15, and 964 cases from the June–July and August–September collection periods were resent on October 15. Collection for resent cases was permitted to continue until the survey completion date of November 30, 2009. In total, a response was obtained for 803 (37.8%) of the resent cases.

7. Other issues

Following collection, inconsistent or unusual responses that were not detected by the computer assisted personal interviewing application were resolved. Edits were developed to be performed after data collection at Statistics Canada's head office, and any remaining inconsistencies were manually corrected. For two specific survey modules, the cognition and physical activities modules, special considerations had to be made in order to produce relevant and accurate results.

7.1 Cognition module

The purpose of the cognition module was to assess the cognitive functioning of respondents as well as to identify gradual changes in cognitive ability due to aging. This module contained four timed tasks:

- A memory test, where the respondent was read a list of 15 words and asked to recall as many of them as possible in 30 seconds. The list of words came from the Rey word list, which is a modified version of the Rey Auditory Verbal Learning Test (Taylor 1959);
- A test of executive functioning, where the respondent was required to name as many animals as possible in one minute;
- The Mental Alternations Test (Teng 1995), which included a sub-task where the respondent was asked to alternate numbers and letters (1A, 2B, 3C, etc.) in 30 seconds; and
- A second recall of the list of words heard in the first memory task.

If permission was given by the respondent, responses to the above tasks were audio recorded. Processing and coding of the recorded responses were done at the head office, and the coded data were integrated with the respondent's survey data.

The cognition module was administered in English and French (the two official languages in Canada) to non-proxy respondents who consented to participate. Due to these limitations, the response rate for the cognition module was significantly lower than the overall survey response rate. Out of 30,865 survey respondents, 25,864 respondents completed at least one of the four timed cognition tasks.

Weighting the responses to the survey respondent total would have resulted in cognition non-respondents having "not stated" values for any incomplete task. In consultation with survey analysts, this was considered an undesirable approach given the importance of the module. Instead, it was decided to calculate a separate cognition weight. This weight was to be used for any analyses involving the cognition module, while the main weight variable was to be used for all non-cognition analyses.

7.2 Physical activities module

The physical activities (PA2) module asked respondents a series of questions about their level of physical activity over the past seven days. The first six questions in the module categorized activities based on intensity levels. The levels were: (1) sitting, (2) walking, (3) light sports or recreational activities, (4) moderate sports or recreational activities, (5) strenuous sports or recreational activities and (6) muscle strength and endurance. The module also collected the number of times respondents had participated in each activity level in the past seven days (frequency), the number of minutes per day that they had

participated in each level (duration) and the specific activities that they had participated in (this consisted of an activity list followed by a write-in option).

During editing, an attempt was made to code the write-in options. It was discovered that a large proportion of responses in the (3) light, (4) moderate, (5) strenuous and (6) muscle strength and endurance questions were recorded as write-ins and that they appeared in the activity list of a different intensity level. It was decided to carry out an editing process to correct for this misreporting.

The editing process reassigned the misreported activities to the intensity level that was deemed to be correct in the following manner. Write-ins were manually reviewed and a file containing the responses in need of recoding was created. The editing process calculated the average frequency and duration associated with each write-in response requiring recoding. This average frequency was moved to the correct intensity level, adding it to the existing frequency in that level and subtracting it from the level from which it was removed. Similarly, the duration being recoded was added proportionally based on frequency to the existing duration in the correct intensity level. The duration in the level from which the activity was removed would not change as it was reported as a per day average, unless removal left no activities at that level (in which case duration would be reset to zero). As the module asked for frequency and duration ranges in each of the intensity levels, the median of each response range was used as a starting point in all calculations.

As an example, consider a record that has a recode moving from (3) light activities to (5) strenuous activities with a corresponding median frequency value of 1.5 days and a corresponding median duration value of 90 minutes per day. There is an existing activity in (5) strenuous activities with a median frequency of 3.5 days and a median duration of 45 minutes per day. The moving frequency value will be added to the existing value to get a new value of 1.5 + 3.5 = 5.0 days. The duration will be combined as shown below, resulting in a new duration of 58.5 minutes per day:

$$D_{New} = \frac{(D_{Moving} * F_{Moving}) + (D_{Existing} * F_{Existing})}{F_{Moving} + F_{Existing}} = \frac{(90 * 1.5) + (45 * 3.5)}{1.5 + 3.5} = 58.5$$

This process was completed independently for each recode so that the duration became the weighted average of all combined durations, where frequency was the weight. The activity list variables for each of the levels were updated appropriately with each movement.

Of the 30,865 respondents, PA2 editing was carried out on 3,994 units (or 12.9% of the total). Among these 3,994 units, there was a significant amount of change at the record level to the frequency and duration variables. This change was analyzed in terms of the Physical Activities Scale for the Elderly (PASE) score, a derived variable that provided an overall measure of a respondent's answers to the entire PA2 module. The PASE score is a copyrighted instrument developed in 1991 by the New England Research Institute (Washburn *et al.* 1993). The results of the analysis are presented in Table 2 and discussed below.

Table 2: Effect of PA2 Module Editing on the PASE Score

	Edited Units		All Units	
	Number	**Average PASE Score**	**Number**	**Average PASE Score**
Before Editing	3,994	159.717	30,865	126.509
After Editing		154.877		125.938

Table 2 shows the calculated effect of the PA2 editing process. For the 3,994 units that took part in the editing, there was an average PASE score decrease of 4.840 or 3.03%. As only 12.9% of records were affected by editing, the effect on the overall average PASE score was diluted to a decrease of 0.571 or 0.45%.

The decrease in PASE score was expected as the editing involved deleting some activities that had been "double counted." For example, it was a common occurrence for a respondent to report an activity under two intensity levels. In this case, it would be kept in the level that was deemed to be correct and would be deleted from the other (the deletion of an activity usually led to a decrease in PASE score). It should be noted that the higher PASE score for edited units compared to total units was also expected as edited units always had a write-in response. This guaranteed an activity in at least one intensity level, whereas many units had no activities at all.

The difference in the average PASE score of all units is relatively negligible since a movement of an activity from one intensity level to another would typically result in an increase to one component of the PASE score and a decrease to the other that would, at least partially, offset each other. However, the main goal of editing was to ensure that the master file was as accurate as possible at both the individual activity level and the intensity level by using all available write-in information to categorize responses.

8. Evaluation

8.1 Response rates

Data from the Canadian Community Health Survey on Healthy Aging were released on May 12, 2010. In total, 41,496 of the 52,010 selected units were in-scope for the survey. Out of these, 33,517 households agreed to participate in the survey, resulting in an overall household-level response rate of 80.8%. Among these responding households, 33,517 individuals (one per household) were selected to participate in the survey, out of which a response was obtained for 30,865 individuals, resulting in an overall person-level response rate of 92.1%. At the national level, this yields a combined (household and person) response rate of 74.4% for the Healthy Aging survey. Table 3 provides household-level response rates by province, Table 4 provides the person-level response rates by province and Table 5 provides person-level response rates by age group.

Table 3: Household-Level Response Rates by Province

Province	In-scope Households	Responding Households	Response Rate
Newfoundland and Labrador	2,516	2,185	86.8%
Prince Edward Island	2,120	1,765	83.3%
Nova Scotia	3,079	2,536	82.4%
New Brunswick	2,812	2,396	85.2%
Quebec	7,065	5,649	80.0%
Ontario	9,046	7,159	79.1%
Manitoba	2,985	2,386	79.9%
Saskatchewan	2,856	2,304	80.7%
Alberta	3,821	3,012	78.8%
British Columbia	5,196	4,125	79.4%
Canada	**41,496**	**33,517**	**80.8%**

Table 4: Person-Level Response Rates by Province

Province	Selected People	Number of Respondents	Response Rate	Combined Response Rate
Nfld. and Labrador	2,185	2,010	92.0%	79.9%
Prince Edward Island	1,765	1,650	93.5%	77.8%
Nova Scotia	2,536	2,282	90.0%	74.1%
New Brunswick	2,396	2,225	92.9%	79.1%
Quebec	5,649	5,217	92.4%	73.8%
Ontario	7,159	6,525	91.1%	72.1%
Manitoba	2,386	2,177	91.2%	72.9%
Saskatchewan	2,304	2,184	94.8%	76.5%
Alberta	3,012	2,735	90.8%	71.6%
British Columbia	4,125	3,860	93.6%	74.3%
Canada	**33,517**	**30,865**	**92.1%**	**74.4%**

Table 5: Person-Level Response Rates by Age Group

Age Group	Selected People	Number of Respondents	Response Rate
45 to 54 years	5,767	5,166	89.6%
55 to 64 years	10,161	9,330	91.8%
65 to 74 years	7,478	6,975	93.3%
75 to 84 years	5,691	5,320	93.5%
85 years and over	4,420	4,074	92.2%
Canada	**33,517**	**30,865**	**92.1%**

8.2 Some results

A study by Ramage-Morin *et al.* (2010) using data from the Healthy Aging survey reported that among people living in private dwellings, 76.2% of Canadians aged 45 to 64 and 55.5% of Canadians aged 65 and older are in good health, which is defined as having good functional health, independence in activities of daily living, positive self-perceived

general health and positive self-perceived mental health. Good health was also seen in the presence of chronic conditions such as high blood pressure, arthritis and back problems, all of which were common among people aged 45 and older.

Ramage-Morin *et al.* (2010) also associated good health with eight modifiable health-promoting factors:

- Smoking status (never smoked daily/quit for 15 or more years)
- Body mass index (not obese)
- Sleeps well
- Fruit/vegetable consumption (five or more times per day)
- Good oral health
- Frequent walker
- Frequent social participation
- Low daily stress

84.0% of people aged 45 to 64 and 90.9% of people aged 65 and older reported four or more positive tendencies with these factors. The majority of people aged 65 and older (52.6%) reported having at least six of these factors.

The likelihood of good health was greater depending on the number of factors that were reported. Figure 1 shows the prevalence of good health given the number of health-promoting factors. People aged 65 and older with five or more health-promoting factors were more likely to be in good health than someone aged 45 to 64 with positive tendencies on two or fewer factors.

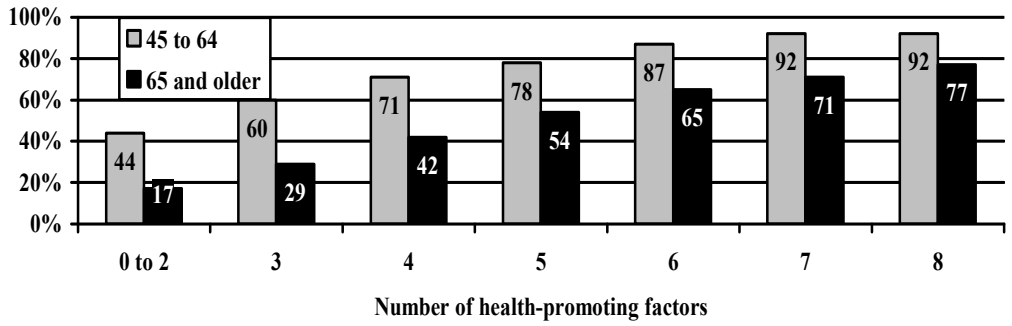

Figure 1: Prevalence of good health by number of health-promoting factors and age group, household population aged 45 and older

9. Conclusion

In this paper, the numerous challenges involved with creating and refining the methodology behind the Healthy Aging survey have been presented. The challenge of finding a suitable frame for the population of interest, Canadians aged 45 and older, was addressed by considering several sampling frames before choosing the 2006 Census of Population. Geographic clusters were created in order to control data collection costs while still achieving the required precision. Stratification and allocation schemes were designed so that older Canadians and those living in rural areas were adequately

represented in the sample. Monitoring during data collection allowed for strategies to be implemented, improving the response rate. Finally, improvements to two modules were made so that survey results remained of a high standard.

The quality guidelines at Statistics Canada (2009) mandate that the quality of information produced by a statistical agency, and its relevance in particular, is of fundamental importance. By using logical and sound methods good results were obtained. The presented approach has hopefully shown that the goal of producing relevant and quality results has been achieved. Further analytical articles based on CCHS–HA and, in particular, the cognition module, are planned.

Acknowledgements

The authors would like to thank Leslie Geran and Cathlin Sarafin for their useful comments and insight. They would also like to thank Michelle Simard and Jack Gambino for their valuable input and recommendations.

References

Bankier, M.D. (1988). Power Allocations: Determining Sample Sizes for Subnational Areas. *The American Statistician*, 42, 3, 174-177.

Park, J. (2010). Health factors and early retirement among older workers. *Perspectives on Labour and Income*, 11, 6, 5-13. Statistics Canada catalogue no. 75-001-X.

Ramage-Morin, P.L., Shields, M. and Martel, L. (2010). Health-promoting factors and good health among Canadians in mid- to late life. *Health Reports*, 21, 3. Statistics Canada catalogue no. 82-003-X.

Statistics Canada (2009). *Statistics Canada Quality Guidelines, Fifth Edition – October 2009.* Statistics Canada catalogue no. 12-539-X.

Statistics Canada (2010a). *Population Projections for Canada, Provinces and Territories.* Statistics Canada catalogue no. 91-520-X.

Statistics Canada (2010b). *Canadian Community Health Survey (CCHS) – Healthy Aging User Guide.* Health Statistics Division. Statistics Canada internal document.

Taylor, E.M. (1959). *Psychological Appraisal of Children with Cerebral Defects.* Cambridge: Harvard University Press.

Teng, E.L. (1995). The Mental Alternations Test (MAT). *The Clinical Neuropsychologist*, 9, 3, 287.

Vijayan, K. (1968). An Exact πps Sampling Scheme – Generalization of a Method of Hanurav. *Journal of the Royal Statistical Society*, Series B, 30, 556-566.

Washburn, R.A., Smith, K.W., Jette, A.M. and Janney, C.A. (1993). The physical activity scale for the elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology*, 46, 2, 153-162.