

# Mode effects in the Canadian Community Health Survey: a Comparison of CAPI and CATI

Martin St-Pierre ([martin.st-pierre@statcan.ca](mailto:martin.st-pierre@statcan.ca)) and Yves Béland ([yves.beland@statcan.ca](mailto:yves.beland@statcan.ca)), Statistics Canada

This article should be cited as: St-Pierre, M. et 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*. Toronto, Canada: American Statistical Association.

**Key words: interview mode effects, CAPI, CATI.**

## 1. Introduction

The Canadian Community Health Survey (CCHS) consists of two cross-sectional surveys conducted over a two-year repeating cycle. The first survey (2001, 2003, 2005, etc.) collects data from over 130,000 households on a range of population health topics and aims to produce reliable estimates at the health region level. The second survey (2002, 2004, 2006, etc.), with a sample size of about 30,000 households, focuses on a particular topic that changes every cycle and aims to produce reliable estimates at the province level (mental health, nutrition, health examination measures, etc.).

The first survey of the first cycle (cycle 1.1), conducted in 2001, made use of multiple sampling frames and data collection modes (Statistics Canada, 2003). In cycle 1.1 the main source for selecting the sample of households was an area probability frame. Field interviewers conducted either personal or telephone interviews using a questionnaire designed for computer-assisted interviewing (CAPI or CATI). The sample was complemented by households selected from either a Random Digit Dialling frame or a list frame of telephone numbers where call centre interviewers conducted CATI interviews with the selected respondents. For operational and budgetary reasons the ratio of area/telephone frame cases changed for the CCHS cycle 2.1 to increase the number of cases completed through CATI. Table 1 shows the change in the sample allocation between the two cycles. It was anticipated that such change in the method of collection would affect the comparability of some key health indicators over the two cycles either by artificially amplifying or masking a real change in behaviours. The percentages in the table below reflect the fact that some area frame units and all telephone frame units are interviewed through CATI.

**Table 1. Sample allocation by frame and mode**

		Cycle 1.1 (2001)	Cycle 2.1 (2003)
Frame	Area	80%	50%
	Telephone	20%	50%
Mode	CAPI	50%	30%
	CATI	50%	70%

A study conducted using the CCHS cycle 1.1 data indicated possible mode effects between CAPI and CATI; this study however had many limitations as some uncontrolled factors distorted the interpretation of the study results (Pierre and Béland, 2002).

In order to better understand the differences caused by the methods of collection (CAPI and CATI) in a large health survey, it was decided to design a special mode study and fully implement it as part of the CCHS cycle 2.1. Although it is understood that many factors could explain differences in survey estimates, it is believed that the results of this study will provide valuable indications to CCHS users on the magnitude of the differences in some key health-related estimates caused by the method of data collection.

This paper presents the results of the mode study. First, the methodology of the study is presented in section 2. It is followed by a summary of the collection procedures. A short description of the processing, weighting and estimation strategy is given in section 4. The results of the mode study are presented in sections 5 and 6 where several univariate and multivariate analyses were performed to assess the presence and the magnitude of the mode effects. A discussion of the results is given in section 7. Finally, a conclusion and some recommendations are provided in last section.

## 2. Methodology of the Study

Due to operational constraints, the mode study was fully embedded in the CCHS cycle 2.1 with minimal modifications to the regular collection procedures. It is important to emphasize that it was not a true experimental design to measure pure mode effects because not all factors were controlled in the design (e.g. interviewers could not be randomized between the two modes of collection). This study however makes use of a split-plot design, i.e., a stratified multi-stage design where the secondary sampling units are randomly assigned to the two mode samples.

### 2.1. Sample Size and Allocation

In order to detect significant differences between point estimates at a certain  $\alpha$ -level, a minimum sample size of 2,500 respondents was targeted for

each mode sample. With such sample sizes and considering the study design effect, a 2%-difference for a 10%-prevalence and a 3%-difference for a 25%-prevalence can be detected at the level  $\alpha=5\%$ .

To facilitate the implementation of the study design with minimal disturbance to the regular CCHS collection procedures it was decided to conduct the study in a limited number of sites (health regions) in Canada. The 11 sites identified for this study provide a good representation of the various regions in Canada (East, Quebec, Ontario, Prairies and British Columbia). Rural health regions with very low density population were not considered for this study for collection cost purposes.

Each mode's sample size was allocated to the study sites proportionally to the CCHS cycle 2.1 sample sizes. Table 2 provides a detailed distribution of the mode study sample by site.

**Table 2 – Mode Study Sample Sizes**

Health Region	CAPI	CATI
St. John's, N.L.	135	100
Cape Breton, N.S.	125	100
Halifax, N.S.	200	150
Chaudière-Appalaches, Qc	230	215
Montréal, Qc	405	390
Niagara, Ont.	235	230
Waterloo, Ont.	235	230
Winnipeg, Man.	320	320
Calgary, Alb.	350	290
Edmonton, Alb.	335	290
South Fraser, B.C.	240	240
<b>Total</b>	<b>2,810</b>	<b>2,555</b>

Extra sample was attributed to CAPI in anticipation of possible telephone interviews (e.g. interviewer must finalize a case over the phone for various reasons); these cases were later excluded. These sample sizes were boosted before data collection to take into account out-of-scope dwellings, vacant dwellings and anticipated nonresponse.

## 2.2. Frame, Selection and Randomization

In the selected sites the CCHS 2.1 used two overlapping sampling frames: an area frame and a list frame of telephone numbers. However and with the objective of eliminating all possible sources of noise during data analysis it was decided to select the mode study sample from one sampling frame only. In order to keep to a minimum the changes to the regular CCHS data collection procedures it was determined

that selecting the sample from the list frame of telephone numbers and assigning the method of collection afterwards would cause less changes in the procedures than selecting from the area frame.

The list frame of telephone numbers used by CCHS cycle 2.1 is created by linking the Canada Phone directory, a commercially available CD-ROM consisting of names, addresses and telephone numbers from telephone directories in Canada, to Statistics Canada internal administrative conversion files to obtain postal codes. Phone numbers with complete addresses are then mapped to health regions to create list frame strata.

As mentioned earlier, the mode study makes use of a stratified two-stage design. The 11 sites represent the study design strata. The first-stage units were the Census Sub-Divisions (CSD) while the telephone numbers were the second-stage units. Within each site, the sample of telephone numbers was selected as follows:

- i. First stage: PPS-selection of CSDs;
- ii. Allocation of the total sample (CAPI + CATI) of a given site to the sampled CSDs proportionally to their sizes;
- iii. Second stage: Random selection of telephone numbers in each CSD.

Once the sample of telephone numbers was selected those cases for which a valid address was not available were excluded from the process and added to the regular CCHS cycle 2.1 CATI sample. Those telephone numbers, which represented approximately 7% of all numbers, would have caused the implementation of severe changes to the procedures for the field interviewers (CAPI method of collection) to perform personal interviews; it was hence decided to exclude them for both mode samples.

Finally and controlling for the CSD within each study site the telephone numbers with a valid address were assigned a method of collection (CAPI or CATI) on a random basis to constitute the two mode samples.

## 3. Data Collection

The data collection for the CCHS cycle 2.1 started in January 2003 and ended in December 2003. The sample units selected from both the area frame and the telephone frame were sent to the field or to the call centres on a monthly basis for a 2-month collection period (there was a one-month overlap between two consecutive collection periods). Two

weeks prior to a collection period, introductory letters describing the importance of participating in the survey were sent to all cases (area and telephone frames) for which a valid mailing address was available.

For the regular area frame cases the field interviewers were 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 was present then the interviewer conducted a personal interview. If not then the interviewer had the choice of coming back at a later date for a personal interview or completing the interview over the phone (in CCHS cycle 2.1, 40% of the area frame cases were completed over the phone).

For the telephone frame cases the call centre interviewers were instructed to assess the status of the phone numbers (specific questions are included in the computer application), list all household members and conduct an interview with the selected individual at that moment or at a later date.

The data collection for the mode study took place between July and early November 2003. For the CAPI mode sample only a subset of field interviewers (experienced and inexperienced) per site were identified to work on the study cases to facilitate the monitoring of the operations. In early July the interviewers received the mode study cases (between 20 and 60) in a separate assignment than their CCHS assignment to clearly identify them as they were instructed to conduct only personal interviews (CAPI). To provide maximum flexibility to the interviewers the collection period for the mode study cases was extended to three months.

The CATI mode sample cases were divided into three and simply added to the CCHS monthly CATI samples (July, August and September) for a two-month collection period. The CATI mode study sample was completely transparent to the call centre interviewers. Those cases were known only by head office staff.

### **3.1. Response Rates**

In total and after removing the out-of-scope units, 3,317 households were selected to participate in the CAPI mode sample. Out of these selected households a response was obtained for 2,788, giving a household-level response rate of 84.1%. Among these responding households 2,788 individuals (one per household) were selected out of which 2,410

responded, giving a person-level response rate of 86.4%. The combined response rate observed for the CAPI mode sample was 72.7%.

For the CATI mode sample, 3,460 in-scope households were selected to participate in study. Out of these selected households a response was obtained for 2,966, giving a household-level response rate of 85.7%. Among these responding households 2,966 individuals (one per household) were selected out of which a response was obtained for 2,598, giving a person-level response rate of 87.6%. The combined response rate observed for the CATI mode sample was 75.1%.

As anticipated, the response rates observed in the mode study (especially for CAPI) are lower than the CCHS cycle 2.1 response rates because the extensive nonresponse follow-up procedures in place for the main survey were not fully implemented for the mode study cases for operational reasons.

## **4. Data Processing, Weighting and Estimation**

As the mode study was fully integrated with the CCHS cycle 2.1 the data collected for the study cases were processed using the CCHS processing system along with the remaining part of the CCHS sample. In addition to the main sampling weight, mode study respondents were assigned a separate and specific sampling weight just for the mode study to fully represent the target population of the 11 sites. The reader should note that the mode study cases were also part of the CCHS cycle 2.1 master data file as well.

Two weighting strategies with various adjustments were processed side-by-side (one for CAPI and one for CATI). Key factors determined the weighting strategy for each mode sample such as:

- use of stratified, multistage design, involving pps-sampling of PSUs and simple random sampling of telephone numbers;
- household-level nonresponse;
- random selection of one person based on household composition;
- person-level nonresponse.

The sampling weights of each mode sample were calibrated using a one-dimensional poststratification of ten age/sex poststrata (i.e. 12-19, 20-29, 30-44, 45-64 and 65+ crossed with the two sexes).

Similarly to the regular CCHS and because of the complexity of the study design, sampling error for the

mode study was calculated using the bootstrap resampling technique with 500 replicates (Rust and Rao, 1996). All results presented in this paper used the mode study sampling weights.

## 5. Univariate Analysis

The main purpose of the mode study was to compare health indicators derived from data collected in-person (CAPI) and those collected over the phone (CATI). This section presents univariate analyses comparing the two modes of collection. First, chi-square tests for association were used to compare the two mode samples in terms of socio-demographic characteristics. All comparisons were performed on weighted distributions and the adjusted chi-square tests for association used a 5% level of significance. Direct comparisons of several health indicators between the two modes are then presented. For these comparisons, Z-tests were applied to see if there was a significant difference between the estimates. Bootstrap weights were used to calculate standard deviations. As the two mode samples were not independent, the standard deviation of the difference between the estimates was calculated by measuring the dispersion of the 500 differences of estimates using the 500 bootstrap replicates. For all health indicators, item nonresponse was excluded from any analysis unless mentioned otherwise. By doing so, it is assumed that item nonresponse is similarly distributed as item response which might not be totally true. It should however be noted that item nonresponse was very low for each mode. A comparison of the household-level and person-level nonrespondents observed in the two mode samples is also presented.

### 5.1. Comparisons of socio-demographic and household characteristics

Although both mode samples are representative of the target population and sampling weights were calibrated to age/sex groupings, differences could still be observed for other socio-demographic or household characteristics. In order to assess those possible differences a series of chi-square tests for association were performed.

The results of the tests can be separated in two groups: the characteristics for which no statistical differences were found between the two mode samples and those for which differences were found. No differences in the distributions were found for the following characteristics: living arrangement, household size, education of respondent, race, immigration and job status. Statistically significant

differences were however found for the following characteristics: marital status, language of interview, highest level of education in the household and household tenure. The main differences can be summarized as follows:

- more single persons in CATI compared to CAPI (31% versus 29%);
- more home owners in CATI (82.7% versus 79.5%);
- more CATI households where the highest level of education was a post-secondary degree (74.4% versus 71%) and;
- more interviews were conducted in another language than English for the CATI sample (27% versus 25.7%).

For the income variables, the item nonresponse was too high to allow for valid comparisons.

### 5.2. Comparisons of health indicators

Statistical Z-tests were performed to determine if the differences were significantly different. Around 70 health indicators for various age/sex domains of interest were looked at and significant differences were found for 15 indicators. Table 3 shows point estimates of selected indicators at the national level (11 sites) by mode.

The most important indicator for which significant differences were found is the obese category of the Body Mass Index (BMI). The CCHS cycle 2.1 collected self-reported height and weight from which a BMI was derived. According to the World Health Organisation, a person is considered obese if his/her BMI is 30 or higher. The obesity rate derived from mode study respondents aged 18 or older is significantly higher for CAPI (17.9%) than for CATI (13.2%). Larger differences were even observed for the 30-44 age grouping (18.1% CAPI and 11.4% CATI) and for men (20.4% and 14.7%).

Another important indicator for which significant differences were found is the physical activity index. The physical activity index is an indicator that shows the amount of leisure-time physical activity done by a person during the last 3 months. It is derived from a series of questions that ask if the respondent has done any of 20 different activities, how many times and for how long. There are significantly more inactive persons in CAPI (42.3%) than with CATI (34.4%).

**Table 3. Comparison of health indicators between CAPI and CATI ( \* = p < 0.05, \*\* = p < 0.01)**

Health indicator	CAPI		CATI		Difference
	%	95% C.I.	%	95% C.I.	%
Obesity (self-reported height and weight)	17.9	15.9-19.9	13.2	11.4-15.1	4.7**
Physical inactivity	42.3	39.5-45.1	34.4	31.8-36.9	7.9**
Current daily or occasional smokers – all ages	23.6	20.7-26.5	21.7	19.8-25.4	1.9
Current daily or occasional smokers – 20 to 29 yrs	37.7	31.4-44.0	28.2	21.7-34.8	9.5*
Alcohol drinker	80.7	78.0-82.5	78.8	76.8-80.8	1.9
At least one chronic condition	69.5	66.5-72.5	68.5	66.2-70.8	1.0
Activity limitation	25.4	22.9-27.8	26.8	24.0-29.5	-1.4
Fair or poor self-rated health	9.3	7.9-10.7	9.9	8.6-11.1	-0.6
Fair or poor self-rated mental health	4.0	2.8-5.2	3.9	2.9-4.9	0.1
Contact with medical doctors in past 12 months	83.5	81.5-85.6	78.4	76.2-80.6	5.1**
Contact with medical specialists in past 12 months	31.1	28.4-33.8	24.9	22.3-27.5	6.2**
Self-reported unmet health care needs	13.9	12.0-15.8	10.7	9.0-12.3	3.2*
Driven a motor vehicle after 2 drinks	13.5	11.3-15.7	7.2	5.1-9.3	6.3**
Ever had sexual intercourse	90.2	88.5-91.9	87.3	85.1-89.5	2.9*

For the smoking indicator (daily or occasional smokers), the rate is 2% higher for CAPI (23.6%) than for CATI (21.7%), but it is not statistically different at the 5% level of significance. However, a significant difference was observed for the 20-29 age group (37.7% for CAPI and 28.2% for CATI). Other results show that the proportion of persons reporting contacts with medical doctors and contacts with medical specialists are higher for the sample interviewed in person. However, the comparisons for contacts with medical doctors broken down by gender shows interesting results where significant differences were found for men (80.3% for CAPI versus 72.5% for CATI) and not for women (86.7% for CAPI versus 84.1% for CATI). As well, significantly more unmet health care needs have been reported for CAPI (13.9%) than for CATI (10.7%).

### 5.3 Comparisons of nonrespondents

Within the CCHS cycle 2.1 and the mode study, total nonresponse could be divided into two categories: household-level and person-level nonresponse. Very little information is known for the 529 CAPI and 494 CATI non-responding households but a comparison of the reasons for not responding shows no major differences between the two modes. For the “no one home/no contact” category the rate for CAPI was 3.6% and 2.1% for CATI. The “refusal” rates are also similar – 8.7% for CAPI versus 10.4% for CATI. Person-level nonresponse is observed when interviewers successfully get through the first part (complete roster with age, sex, marital status and highest level of education of all members) but not the second part, the actual CCHS interview with the

selected respondent. Table 4 compares the age group distributions of the nonrespondents (person-level) observed in CAPI and CATI. It is interesting to note the differences at the two ends of the age groups. A response from elderly persons (65 and up) is much more difficult to obtain over the phone (13.9% nonresponse) than in person (8.9%) while the opposite is observed for the younger age group (12-19). Although the variable “age” is used in the creation of the response propensity classes for the person-level nonresponse weighting adjustment, the nonresponse bias could be non-negligible for some characteristics. One could think that elderly persons with a physical condition might have difficulty to get to the phone. The same could be said with teenagers where the more physically active ones could be home less often and hence less available for a personal interview. This would however require further research.

**Table 4. Person-level Nonresponse Rate (%)**

Mode	Total	12-19	20-29	30-44	45-64	65+
CAPI	13.6	17.6	15.7	15.1	12.4	8.9
CATI	12.4	11.9	16.9	12.0	10.1	13.9

## 6. Multivariate Analyses

To better understand the differences and to ensure that the mode effects found in the indicators comparisons are not simply due to discrepancies in the socio-demographic characteristics between the two mode samples, a series of multiple logistic regressions were performed. This analysis evaluates the effect of the mode of collection on the prevalence of several health indicators when controlling for the

socio-demographic and household variables. The mode effect is treated as a confounded variable in the model. The socio-demographic variables are other confounded variables. Interaction terms between the mode of collection and the socio-demographic variables were all tested in the model.

For selected health indicators, table 5 shows the odds of having the health condition or the health determinant when interviewed by telephone in comparison of when interviewed in person.

The first result presented concerns the smoking indicator. Results in section 5.2 did not show a significant mode effect at the national level for that variable. This analysis shows that for white persons between 12 and 29 years old, being interviewed by telephone makes their odds of reporting a current daily or occasional smokers about 1.8 times ( $1/0.56 = 1.79$ ) less than if interviewed in person (significantly different at the 1% level). For white persons 30 years old and over, the odds are the same (1.00) for CATI and CAPI. For non-white persons, being interviewed by telephone makes their odds of reporting a current daily or occasional smoker about 1.5 times (1.49) more than if interviewed in person, but it is not significant at the 5% level.

As presented in section 5.2, being interviewed by telephone makes the odds of reporting obese lower than if interviewed in person. These odds are even lower in Alberta (0.48); elsewhere in Canada the odds are 0.79. For the physical activity index (inactive), no interaction was found between the mode of collection and the socio-demographic variables. Overall, being interviewed by telephone makes their odds of reporting inactive about 1.5 times ( $1/0.65 = 1.54$ ) less than if interviewed in person.

For the alcohol use indicators, ethnicity, education and age group are characteristics for which mode effect is found. White non-immigrant persons are less likely to describe themselves as alcohol drinker when interviewed by telephone (odds = 0.7), whereas the opposite is observed for non-white or immigrant persons (odds = 1.71). Similarly, for non-white persons, being interviewed by telephone makes their odds of reporting to have had 5 or more drinks in one occasion at least once a month about 2.5 times more than if interviewed in person. The opposite mode effect is found for white persons in the lowest or the lower income adequacy category (odds=0.45).

For the drinking and driving characteristics, a mode effect is found in the 20 to 44 age group. For these

persons, being interviewed by telephone makes their odds of reporting drinking and driving about 3.4 times ( $1/0.29$ ) less than if interviewed in person.

Another result shows that the persons not in the highest income adequacy category and without a post-secondary degree are less likely to report unmet health care needs when interviewed by telephone.

## 7. Interpretation of the results

The results of the mode study are quite diverse. Nearly no differences were found between CAPI and CATI in the point estimates for the vast majority of health indicators measured by CCHS such as tobacco use (all ages), chronic conditions, activity limitations, fruit and vegetable consumption and others. This means that the comparability of the health indicators over the first two cycles of CCHS is not affected by the increased number of CATI in the second cycle.

Significant differences were however found between CAPI and CATI for some health indicators. Among others, self-reported height and weight, physical activity index, contact with medical doctors and self-reported unmet health care needs are certainly the most notable ones. Although the multivariate analysis somewhat attenuated the impact of the mode effects when socio-demographic characteristics are considered, it is believed that any comparison of the above indicators over the two cycles should take into consideration the increased number of CATI in the second cycle. It is important to mention that other methodological (sample sizes, reference period, questionnaire, etc.) and contextual (changes in standards, true change, etc.) aspects should, as well, always be taken into consideration in any comparison of survey indicators over time.

Extensive literature exists on comparisons between personal and telephone interview techniques and a great deal of inconsistencies in the results is certainly noticeable as these studies report varying magnitude of mode effects. Scherpenzeel (2001) suggests that the inconsistency among results is probably caused by differences in the design of the studies. The mode study conducted as part of the CCHS cycle 2.1 is no exception as no comparable studies could be found. There is however unanimity on the presence of mode effects for some variables and the non-negligible biases on survey estimates.

**Table 5. Odds ratios of the health condition for CATI versus CAPI ( \* =  $p < 0.05$ , \*\* =  $p < 0.01$ )**

Health indicator	Factor	Odds ratio
Smoking	White 12-29	0.56**
	White 30+	1.00
Obesity (self-reported)	Non-White	1.49
	Alberta	0.48**
Physical inactivity	Elsewhere	0.79*
	All	0.65**
Influenza immunization	12-15	4.48**
	16-19	1.78
	20+	1.10
Alcohol drinker	White non-immigrant	0.70**
	Non-white or immigrant	1.71**
5 or more drinks on one occasion at	White and lowest or lower middle income	0.45*
	White and highest or higher middle income	0.97
	Non-white	2.45*
Unmet needs (self-reported)	Highest income adequacy	1.11
	Not highest income adequacy but with post-secondary degree	0.81
	Not highest income adequacy and no post-secondary degree	0.46**
Drinking and driving	12-19	1.23
	20-44	0.29**
	45-64	0.97
	65+	0.60
Ever had sexual intercourse	Female 15-24	0.43*
	Others	1.02

The authors of this paper think that the differences found in the mode study of the Canadian Community Health Survey between CAPI and CATI are mainly caused by two confounding factors: social desirability and interviewer variability. The widely documented social desirability response bias is generated by people's attempts to construct favourable images of themselves in the eyes of others. It could occur at different levels and for different topics for both CAPI and CATI and it is very difficult to quantify the magnitude of the measurement biases due to the absence of "gold standards" for many variables. Moreover the magnitude of the bias would differ based of socio-demographic profiles and it could even vary in time. Among all health indicators evaluated in this study, self-reported height and weight are good examples of variables for which the magnitude of the social desirability response biases differ between CAPI and CATI. Preliminary data of the 2004 Canadian Nutrition Survey conducted by Statistics Canada where exact measures of height and weight are collected on a large sample suggest that the obesity rate among Canadians of all ages is significantly higher than those calculated using the self-reported measures of the CCHS cycle 2.1 mode study (CAPI

and CATI). Clearly the measurement bias is larger in CATI than in CAPI but they are both far from the "gold standard" derived from the nutrition survey. The reader should note that the results of the 2004 Canadian Nutrition Survey will be available in the fall of 2005.

The interviewer variability is the term used to describe the errors that are attributable to interviewers. Interviewer variability is inevitable in large surveys conducted by National Statistical Organisations. At Statistics Canada, the field interviewing staff is composed of more than 650 interviewers and 250 interviewers work in the call centres. Despite all efforts to standardize training procedures among all interviewers some aspects of the work environments (*e.g.* supervision) of the two collection methods are simply so different that it is reasonable to believe that interviewers' behaviours could differ from one to the other and hence interviewer variability biases could be introduced. For the mode study, additional information provided by the computer application systems (CAPI and CATI) such as time length of each question revealed interesting findings. The physical activity module of the CCHS questionnaire from which the physical

activity index is derived took significantly less time to conduct in CAPI than in CATI suggesting that some activities (from the list of 20 activities read by the interviewers) might not have been clearly mentioned to some CAPI respondents for various reasons. In parallel, the quality control procedures implemented in the call centres have not detected such behaviours from the CATI interviewers. The authors believe that the interviewer variability explains a large part of the differences observed in the mode study for the physical activity index but the absence of a gold standard for this variable does not allow for an assessment of the real measurement bias (CAPI or CATI).

## 8. Conclusion

The mode study was fully integrated as part of the CCHS cycle 2.1 to better understand potential differences caused by the two methods of collection used in the CCHS – CAPI and CATI – on survey estimates. It was anticipated that the increased number of CATI interviews in cycle 2.1 compared to cycle 1.1 would affect the comparability of some key health indicators over the two cycles either by artificially amplifying or masking a real change in behaviours.

The mode study used a split-plot design with a unique sample frame where the secondary sampling units were randomly assigned to either CAPI or CATI. The study was conducted between July and November 2003 in 11 sites selected to provide a good representation of each region in Canada. Acceptable response rates were observed for each mode of collection and although minor differences were detected in the socio-demographic profiles the two mode samples are representative of the target population and are comparable. Special sampling weights were computed and calibrated to ten age/sex post-strata for each mode sample. It is important to mention that it was not a true experimental design to assess pure mode effect. However the mode study was designed to allow for valid comparisons between CAPI and CATI collection methods as conducted by Statistics Canada.

The results of the mode study are very useful to better understand the differences between CAPI and CATI and especially the impact of increased CATI in cycle 2.1 compared to cycle 1.1. As well and in light of the observed results, a series of recommendations has been made for future cycles of CCHS. First it was decided to implement the same cycle 2.1 sample design (area/telephone frames and CAPI/CATI ratios) for CCHS cycle 3.1 scheduled for January

2005. Starting in CCHS cycle 3.1, exact height and weight will be collected on a subsample of individuals to allow for national estimates of BMI categories for specific age/sex groupings. Also, interviewers' procedures will be reinforced to standardize even more collection procedures among the two collection methods.

These improvements should hence improve the quality of CCHS data and provide a solid basis to policy makers and health care professionals to better track changes over time and take appropriate actions to address the various issues around the health of Canadians.

## 9. Acknowledgments

The authors would like to thank all their colleagues at Statistics Canada who participated in the development and realisation of this study. They are also grateful to Vincent Dale, Johane Dufour and Jean-Louis Tambay for their insightful comments.

## 10. References

- Pierre, F. and Béland, Y. (2002). *Étude sur quelques erreurs de réponse dans le cadre de l'Enquête sur la santé dans les collectivités canadiennes*. 2002 Proceedings of the Survey Methods Section, Statistical Society of Canada.
- Rust, K.F. and Rao, J.N.K (1996). "Variance estimation for complex surveys using replication techniques", *Statistical Methods in Medical Research*, 5, p. 281-310.
- Scherpenzeel, A. (2001). *Mode effects in panel surveys: A comparison of CAPI and CATI*. Bases statistiques et vues d'ensemble. Neuchâtel: Bundesamt, für Statistik, Office fédéral de la statistique (<http://www.unine.ch/psm>).
- Statistique Canada (2003). *CCHS Cycle 1.1 2000-2001 Public Use Microdata Files*. Catalogue no. 82M0013GPE.