Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2006 : Methodological Issues in Measuring Population Health

2006

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada's Symposium 2006
Methodological Issues in Measuring Population Health

Accounting for a Significant Methodological Change in Analyzing Canadian Community Health Survey Data

Luc Côté and Robert Courtemanche[1]

Abstract

Statistics Canada's Canadian Community Health Survey uses two sample frames and two data collection methods. In cycle 2.1, a change was made in sample allocation between the two frames. A study of the collection method effect by Statistics Canada revealed comparability problems between cycles 1.1 and 2.1. In contrast, the Institut de la statistique du Québec took a comprehensive look at the changes, and classified 178 variables as "comparable" or "non-comparable". It made recommendations to Quebec users concerning chronological and interregional comparisons.

KEYWORDS: Comparability over time; cross-sectional estimates; interregional comparisons; collection method; multiple frames; sample allocation

1. Introduction

The results of Statistics Canada's study on the effect of the collection methods (telephone and personal interview) (St-Pierre and Béland, 2004) indicated that there were differences in a number of the survey's variables, which in turn pointed to data comparability problems between cycles 1.1 and 2.1 of the Canadian Community Health Survey (CCHS). As a result, Statistics Canada recommended caution in using the data. This quickly led to questions about comparability: Can we compare the data of the CCHS's first two cycles or not? Which variables can we compare? What can we say about the ones whose estimates we cannot compare?

With the aim of answering these questions – and not of finding the cause of the differences in the two cycles' data – the Institut de la statistique du Québec (ISQ) (Côté, Courtemanche and Caron, 2005) examined the main methodological change made in CCHS cycle 2.1, a change in the sample allocation between the two frames.

Section 2 provides a brief description of the CCHS's sampling plan and Statistics Canada's study of the collection methods, and outlines the objective of the ISQ's study. Section 3 focuses on the statistical methodology used to achieve the objective. Some of the results are presented, and recommendations are made to users concerning the study's objective (section 4) and interregional comparisons (section 5). Section 6 presents the article's conclusions.

2. Background and Objective of the ISQ Study

Statistics Canada conducted the Canadian Community Health Survey (CCHS) in 2000-2001. It was a regional cross-sectional survey on various topics related to the health of Canadians. It was carried out again in 2003. The two surveys are known as cycles 1.1 and 2.1 of the CCHS (Statistics Canada, 2003 and 2005).

In each cycle, precise statistics were expected for about 130 health regions in Canada. In Quebec, there were 16 health regions in cycle 1.1 and 17 in cycle 2.1.

[1]Luc Côté and Robert Courtemanche, Institut de la statistique du Québec, Direction de la méthodologie, de la démographie et des enquêtes spéciales, 200 Sainte-Foy, 3rd Floor, Québec City, Quebec, Canada, G1R 5T4 luc.cote@stat.gouv.qc.ca and robert.courtemanche@stat.gouv.qc.ca.

For most regions, two sample frames were used: an area frame composed of geographic units used in the Canadian census, and a telephone frame consisting of telephone numbers listed in directories. For some parts of Canada, however, the area frame did not provide sufficient coverage, and a frame consisting exclusively of randomly generated telephone numbers was used.

In regions where two frames were used, one sample was selected independently from each frame. In all, more than 130,000 people responded to the survey. In Quebec, questionnaires were completed by some 22,000 respondents in cycle 1.1 and just over 27,000 in cycle 2.1.

Two collection methods were used. For the selected persons in households sampled from the telephone frame, a computer-assisted telephone interview (CATI) was administered. For persons in households selected at random from the area frame, a computer-assisted personal interview (CAPI) was conducted in most cases, although Statistics Canada had planned to do some telephone interviews.

As shown in Table 1, about 95% of the weighted sample for Quebec was from the area frame in cycle 1.1. In cycle 2.1, the area frame accounted for only 44%. The table also reveals that the proportion of the area-frame sample that was interviewed in person declined from 74% in cycle 1.1 to less than 61% in cycle 2.1.

Table 1
Allocation of the weighted sample by collection method and sample frame, Quebec, CCHS, cycles 1.1 and 2.1

Cycle 1.1

Sample frame	**Collection method %**		**Frame total[1] %**
	CAPI	CATI	
Area	73.8	26.2	**95.2**
Telephone	0.0	100.0	**4.8**
Method total[2]	**70.1**	**29.9**	**100.0**

Cycle 2.1

Sample frame	**Collection method %**		**Frame total[1] %**
	CAPI	CATI	
Area	60.6	39.4	**44.3**
Telephone	0.0	100.0	**55.7**
Method total[2]	**26.8**	**73.2**	**100.0**

1. Proportion of the weighted sample selected from the area frame or the telephone frame.
2. Proportion of the weighted sample interviewed in person (CAPI) or by telephone (CATI).
Sources: Statistics Canada, data sharing files for CCHS cycles 1.1 and 2.1 (2000-2001 and 2003)

As a result of this significant change in the sample allocation, Statistics Canada carried out a study that was incorporated into the collection process for CCHS cycle 2.1. Its aims were to examine exactly what effect the collection method had on the estimates, and to determine whether the observed differences were solely due to the collection method.

The technique used was based on a split-plot approach, which involved taking a subsample of CCHS cycle 2.1 households from the telephone frame only. A total of 11 health regions in Canada, corresponding to the study's strata, were covered. The primary sampling units, census subdivisions (CSDs), were selected with probability proportional to size. The total sample was allocated proportionally to the size of the selected CSDs, and a sample of listed telephone numbers was selected for each CSD. A collection method (CAPI or CATI) was assigned at random to each telephone number, which generated two independent regional subsamples. For more details, see the article by St-Pierre and Béland (2004).

The study identified differences in the data collected in personal interviews and in telephone interviews for a number of key variables, which causes a comparability problem between cycles 1.1 and 2.1. Statistics Canada recommended that the sample allocation for each collection method and sample frame be kept unchanged in future cycles. It also recommended that users exercise caution in interpreting comparisons between the two cycles.

The change in sample allocation (area frame and telephone frame) between CCHS cycles 1.1 and 2.1 had the greatest impact in three areas that might affect the two cycles' comparability. First, coverage was not the same: the area frame provides better coverage than a frame consisting of listed telephone numbers. Second, better response

rates were obtained with the area frame (83% in Quebec and Canada) than with the telephone frame (75% in Quebec and 78% in Canada), which points to better data quality (Statistics Canada, 2005, p. 38). Third, the collection method varies with the frame: data for the area frame sample were collected mostly through personal interviews, while data for the telephone frame sample were gathered exclusively through telephone interviewing, as shown in Table 1.

The aims of the ISQ study were to test the change in the sample allocation and thus determine the overall effect that making changes in all three areas had on the estimates produced for cycles 1.1 and 2.1. This was also the intent of Groves *et al*. (2004, p. 148). The ISQ study was not intended to find the causes of the differences between the CCHS's first two cycles but rather to determine whether the samples from the two frames yielded comparable statistics, which is different from the goal of Statistics Canada's study. In so doing, we were able to answer the following questions: Can we compare the data from the CCHS's first two cycles or not? Which variables can we compare? What can we say about the ones we cannot compare? In short, what recommendations can the ISQ make for Quebec users?

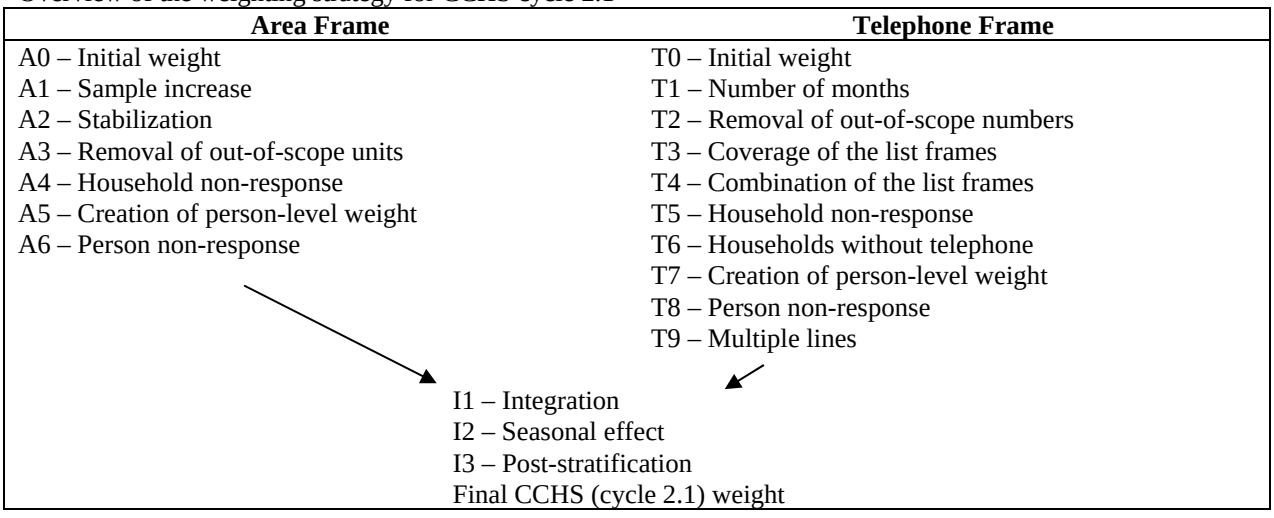
3. Methodology

As noted earlier, the CCHS sample from each frame was selected independently. Cycle 2.1 was ideal for our study, because the regional sample from each frame was sufficiently large, in contrast to cycle 1.1, in which the regional telephone sample was small or non-existent in a number of health regions.

The study was based on the data sharing file produced for Quebec. The file contains data for CCHS respondents who agreed to allow the information they provided to be shared (in Quebec, about 93% consented). The study covered 15 of the 17 regions. The sample for the two excluded regions was drawn entirely from one frame. Those two regions contributed just 0.4% of Quebec's data.

To compare the area frame data with the telephone frame data, we reweighted the cycles independently for each frame. In this regard, it is useful to understand how Statistics Canada carried out the original weighting (Figure 1). First, the steps that differentiated the two frames – the probability of selection and the non-response adjustments at the household and individual levels (steps A0 to A6 for the area frame; steps T0 to T9 for the telephone frame) – were processed independently. Second, the two weighting systems were integrated using a linear combination (step I1). Third, two calibration steps (I2 and I3) were performed to produce the final weight for cycle 2.1. The cycle 2.1 bootstrap weights used to estimate precision were obtained in the same way. The cycle 1.1 weights and the bootstrap weights were constructed in a similar manner.

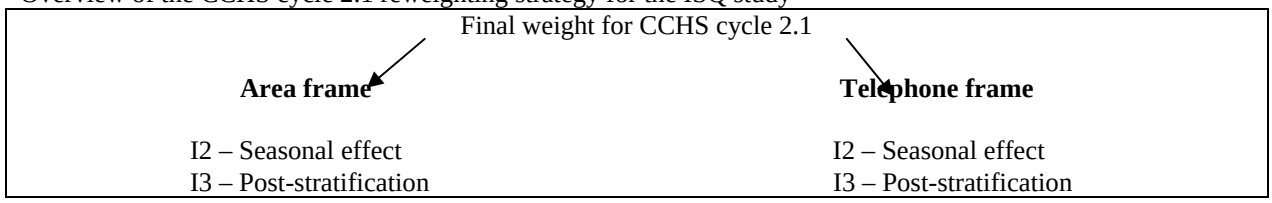
Figure 1
Overview of the weighting strategy for CCHS cycle 2.1



Source: *User Guide for the Public Use Microdata File*, CCHS Cycle 2.1, Diagram A, Statistics Canada, 2005, p. 23.

For reweighting, we took the final weight for cycle 2.1 and decomposed it for each frame (Figure 2). Then we redid steps I2 and I3 separately for each frame. This procedure was repeated for the cycle 2.1 bootstrap weights to produce bootstrap weights for the area frame portion and bootstrap weights for the telephone frame portion. For cycle 1.1, only the area frame weights and the area frame bootstrap weights were redone. Following this, we were able to produce an estimate for the area sample and a separate estimate for the telephone sample for cycle 2.1.

Figure 2
Overview of the CCHS cycle 2.1 reweighting strategy for the ISQ study



Chi-square tests with Sattherthwaite's adjustment were constructed to determine whether there was a correlation between a health variable (for example, consulting a physician) and the sample frame (area or telephone) in cycle 2.1. If the test was not significant at the 5% level, we concluded that in cycle 2.1, the data from the area sample were not significantly different from the data from the telephone sample, and therefore that users could compare the two cycles in the usual manner. Conversely, if the test was significant, it meant that the data from the area sample and the data from the telephone sample did not measure the same phenomenon and that when this result was transposed in time, the two cycles could not be compared in the usual way. The health variables were characterized as "not directly comparable". In those cases, a proportion difference test was performed between cycles 1.1 and 2.1 for the statistic generated with the area sample only, as the telephone sample was small or non-existent in a number of regions in cycle 1.1.

For all those tests, bootstrap weights were used to control for the sampling plan. All the tests were performed with SUDAAN (Research Triangle Institute, 2001).

4. Chronological Comparisons: Results and Recommendations

4.1 Results

A total of 178 "similar" variables (identical wording and response choices in cycles 1.1 and 2.1) were analyzed. For 110 of them, no significant difference was detected. The other 68 variables were declared to be "not directly comparable". The most seriously affected variables were the ones associated with physical activity. We found little difference for the variables associated with health problems.

For the "not directly comparable" variables, Quebec users were informed of the observed direction of the change between cycles 1.1 and 2.1 (significant increase, significant decrease, or non-significant difference). Those users, of course, do not have access to the data sharing file. The direction of change was determined with a proportion difference test that used only area frame estimates.

Table 2 presents the results for the "Number of consultations with a physician (general practitioner or specialist) in the last 12 months" variable. In cycle 2.1, a difference was detected between the area frame data and the telephone frame data. If we consider only the area frame results in our analysis, we find that there is no significant difference between cycles 1.1 and 2.1, since the 95% confidence interval includes the zero point. We would not have reached the same conclusion if we had used the data from both frames combined. We would have concluded wrongly that the number of consultations declined, which is implausible in view of the aging of Quebec's population.

Table 2
First example of results for a "not directly comparable" variable: Number of consultations with a physician (general practitioner or specialist) in the last 12 moths, Quebec, age 12 and over, by frame and cycle, CCHS

Data source	Cycle 2.1	Cycle 1.1	95% confidence interval
Combined	75.9%	77.9%	[-3.3%; -0.8%]
Area frame	78.7%	78.2%	[-1.0%; 2.0%]
Telephone frame	74.0%	-	-

The second example concerns the intensity of physical exertion during recreational activity in the past three months. We were interested in the proportion of inactive people. As shown in Table 3, when the two frames are combined, the proportion of inactive people declines from cycle 1.1 to cycle 2.1. That is what we would conclude if we did not take the methodological change into account. If we used only the area frame data, we would also conclude that the proportion decreased, since the zero point lies just outside the 95% confidence interval. The impact of the methodological change is reflected in the size of the decline. When the combined frame is used, the decrease is an estimated 7%, and when the area frame is used, it is closer to 2%. This example clearly shows that the combined frame overestimates the observed decline between cycles 1.1 and 2.1.

Table 3
Second example of results for a "not directly comparable" variable: Physical exertion during recreational activity in the past three months ("inactive" category), Quebec, age 12 and over, by frame and cycle, CCHS

Data source	Cycle 2.1	Cycle 1.1	95% confidence interval
Combined	51.6%	58.9%	[-8.5%; -5.5%]
Area frame	57.5%	59.3%	[-3.6%; -0.01%]
Telephone frame	46.9%	-	-

4.2 Recommendations

For chronological analysis (between the two cycles), we made recommendations to guide Quebec users in their research. It is worth noting that unlike the ISQ, those users only have access to the public use microdata file, to which confidentiality protection measures have been applied. For example, users do not have access to the bootstrap weights.

For the 110 "comparable" variables, the ISQ recommended that Quebec users follow the usual procedures and construct proportion difference tests for the two cycles using the data from the combined frame, i.e., the data in the data files. For the 68 "not directly comparable" variables, only the sample from the area frame is to be used for Quebec comparisons. Since the users do not have access to the statistics computed from the area sample, we provided them with the direction of change (increase, decrease, no significant difference), but not the size of the change. No recommendation was made concerning regional chronological analyses, since differences cannot be detected at the regional level.

5. Interregional Comparisons in Cycle 2.1: Methodology, Results and Recommendations

The differences we observed in cycle 2.1 between the area frame data and the telephone frame data (see Tables 3 and 4) also raised questions about the impact that those differences might have on interregional comparisons in cycle 2.1, i.e., comparisons between the estimates for one region and estimates for all other regions in Quebec.

In section 2, we noted that in cycle 2.1, the area frame accounted for 44% of the weighted sample in Quebec. What is the proportion at the regional level? As shown in Table 4, it is about 25% in three of the 15 regions (Capitale-Nationale, Outaouais and Laurentides). Those regions purchased an additional sample, which was drawn exclusively from the telephone frame. This difference in sample allocation between the two frames could affect comparisons between each of the three regions and the rest of Quebec, since each one's sample allocation is different from Quebec's. There is also the case of the Montréal-Centre and Montérégie regions, which have a large sample because they are the most populous regions in Quebec. Even though their sample allocations are similar to

that of Quebec as a whole, statistical differences with other regions may be detected because of their statistical power.

Table 4
Proportion of the weighted sample derived from the area frame, by health region, Quebec, CCHS, cycle 2.1

Health Region	**%**
01 Bas-Saint-Laurent	50.4
02 Saguenay–Lac-Saint-Jean	45.3
03 Capitale-Nationale	24.3
04 Mauricie–Centre-du-Québec	39.0
05 Estrie	45.2
06 Montréal-Centre	59.2
07 Outaouais	28.1
08 Abitibi-Témiscamingue	48.3
09 Côte-Nord	43.4
11 Gaspésie–Îles-de-la-Madeleine	57.6
12 Chaudière-Appalaches	48.7
13 Laval	51.6
14 Lanaudière	52.3
15 Laurentides	24.3
16 Montérégie	38.1
Quebec as a whole	**44.3**

Source: Statistics Canada, data sharing file for CCHS cycle 2.1, 2003.

5.1 Methodology

We decided to use an approach based on regions and health variables to determine whether the sample allocation between the two sample frames might affect the results of comparisons between a particular region and all other regions in Quebec. To that end, we used a logistic regression model to test whether a significant difference detected between a particular region and all other regions combined persisted when the "frame" variable was added to the model. The significance level was set at 5%. We examined 23 of the 68 "not directly comparable" variables. We selected some variables for which the difference between the area frame data and the telephone frame data was large, some for which it was medium-sized, and some for which it was small.

For each of the 23 health variables selected, we constructed 15 models. This first set of models did not include "region" as an independent variable. For each of the 345 models, we tested whether the variable was significant at the 5% level.

$$\text{Logit (health variable)} = \beta_0 + \beta_1 \times \text{region} + \varepsilon$$

where region = 1 if the respondent lives in the region in question
= 0 if the respondent lives in another part of Quebec

We constructed a second set of models for each of the 23 health variables by adding "frame" as an independent variable. After adding that control variable to the model, we tested whether the "region" variable was significant at the 5% level.

$$\text{Logit (health variable)} = \beta_0^* + \beta_1^* \times \text{region} + \beta_2^* \times \text{frame} + \varepsilon$$

5.2 Results and recommendations

On the basis of the results obtained with the models, we divided the regions into two classes: one composed of the three regions that purchased additional sample and the two most populous regions, and the other composed of the remaining 10 regions. This distinction is also reflected in the recommendations.

Results for the 10 regions

Controlling for the frame's effect in the logistic regression model had a slight effect on the observed levels for the "region" variable. There were certainly some borderline cases (observed level of about 0.04); some differences between the region in question and all other Quebec regions became significant when the frame was taken into account, while other differences went from significant to non-significant. In view of those results, we concluded that adding the frame to the model had little impact.

Recommendation for the 10 regions

We recommend no adjustment. Quebec users can follow the usual procedure in constructing tests of proportion differences between one particular region and all other regions combined, for both "comparable" and "not directly comparable" variables.

Results for the three regions that purchased additional sample and the two most populous regions

We observed changes in the observed level for the "region" variable depending on the presence or absence of the "frame" control variable. The changes are shown in Table 5. We constructed 115 models that included only the "region" variable (23 health variables multiplied by five regions). For 42 of the 115 models, the "region" variable had an observed level of less than 5%. We also found that in 26 of the 42 cases in which a significant difference was detected, the observed level was less than 1%. When the "frame" control variable was added, the "region" variable went from significant to non-significant at the 5% level for 13 of the 42 models. When we controlled for the frame in the 26 cases with an observed level of less than 1%, four differences became non-significant.

Table 5
Results of comparisons between the data for one specific region and the data for all other Quebec regions combined, by the absence or presence of the "frame" variable, regions that purchased additional sample (Capitale-Nationale, Outaouais and Laurentides) and the most populous regions in Quebec (Montréal-Centre and Montérégie), CCHS, cycle 2.1

Observed level for the "region" variable when the "frame" variable is not in the model	**Models for which the "region" variable is significant (models without the "frame" variable)**	**Models for which the "region" variable becomes non-significant at the 5% level when the "frame" variable is included**
	n	n
Observed level < 5%	42	13
1% ≤ observed level < 5%	16	9
Observed level < 1%	26	4

The above analysis demonstrates that the sample allocation between the two frames has an effect on the tests of differences between one specific region and the regions that purchased additional sample or the two most populous regions in Quebec. Consequently, it is important to take the frame into account in this type of comparison. However, users were unable to include the "frame" variable in regression models because they had no access to the file containing the bootstrap weights or to the "reweighted" weights. We therefore needed to suggest easy-to-follow rules that took their data environment into account.

Recommendation for the three regions that purchased additional sample and the two most populous regions

With regard to the regions that purchased additional sample (Capitale-Nationale, Outaouais and Laurentides) and the two most populous regions in Quebec (Montréal-Centre and Montérégie), the ISQ recommended that differences with an observed level of more than 1% not be discussed for "not directly comparable" variables (models without the "frame" variable) to ensure that the error would not be more than 5% if the "frame" variable were taken into account.

With regard to the "comparable" variables, Quebec users can follow the usual procedures in constructing tests of proportion differences between one particular region and all other Quebec regions and using a significance level of 5%.

6. Conclusion

Following the change in the sample allocation in cycle 2.1 of the Canadian Community Health Survey, the Institut de la statistique du Québec decided to examine its effect on comparisons with cycle 1.1. The ISQ's aim was to provide users with guidelines to assist them in their analyses. This led to the preparation of a list of "comparable" variables and a list of "not directly comparable" variables. In addition, recommendations were made to guide Quebec users in their analyses of chronological and interregional comparisons.

For cycle 3.1 of the CCHS, Statistics Canada planned to stabilize the sample allocation between the frames for Quebec as a whole. However, three regions in Quebec purchased an additional sample in cycle 3.1. Those regions will not be able to use the usual methods in comparing their data.

Acknowledgements

The authors are grateful to Bernard Caron, a co-op student at the Institut de la statistique du Québec in the summer of 2005, for producing the large number of analytical tables, and to Rosanna Baraldi, Monique Bordeleau, Carole Daveluy and Claire Robitaille, of the Institut de la statistique du Québec, for their insightful comments, which led to improvements in the recommendations.

References

Côté, L., R. Courtemanche and B. Caron, with the collaboration of C. Daveluy, C. Robitaille, R. Baraldi et M. Bordeleau (2005), *Comparabilité entre les cycles 1.1 et 2.1 de l'Enquête sur la santé dans les collectivités canadiennes : impact du changement apporté à la répartition de l'échantillon selon la base de sondage*, Québec, Institut de la statistique du Québec, 69 p.

Groves, R. M., F. J. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer and R. Tourangeau (2004), *Survey Methodology*, New York, John Wiley & Sons.

Research Triangle Institute (2001), *SUDAAN User's Manual, Release 8.0*, Research Triangle Park, N.C., Research Triangle Institute.

St-Pierre, M., and Y. Béland (2004), « Mode Effects in the Canadian Community Health Survey. A Comparison of CAPI and CATI », *Proceedings of the Survey Research Methods Section*, American Statistical Association.

Statistics Canada (2003), *CCHS Cycle 1.1, 2000-2001, Public Use Microdata File*, Catalogue Number 82M0013GPE.

Statistics Canada (2005), *CCHS Cycle 2.1, 2000-2001, Public Use Microdata File*, Catalogue Number 82M0013GPF.