Quality measures
Canadian Community Health Survey (CCHS)

(Survey numbers 3226 and 5015)

The CCHS two-year collection cycle comprises a general population health survey at the health region level in the first year with a total sample of approximately 130,000 respondents and a provincial level survey on a specific focussed health topic in the second year with a total base sample of approximately 30,000 respondents.

Over the two-year period covered by this report, collection and processing for Cycle 1.1 (*2000/2001*) of the CCHS were completed. Data was disseminated on May 8, 2002 through an article in the *Daily* and on CANSIM and other derived products on the Statistics Canada web site. Share files were made available to the provincial and territorial ministries of health, Health Canada and l'Institut de la statistique du Québec. The master file was also made available in the Research Data Centres in the following weeks. The Public-Use Microdata File was released in January 2003.

Collection for Cycle 1.2 on mental health and well-being (2002) was completed in December 2002. The results were released September 3, 2003, much in the same was as Cycle 1.1 results were released (*i.e. in* the Daily, CANSIM, research Data Centres).

Coverage

Cycle 1.1 used a mix of the Labour Force Survey (LFS) area frame, a Random Digit Dialling frame (RDD) and a list frame of telephone numbers for its sample. The target population was Canadians 12 years of age or older who are living in private dwellings. Excluded from the area frame were residents of institutions, full-time members of the Canadian armed forces, residents of Indian reserves and of Crown lands, and residents of a few remote areas. The telephone frames excluded people who do not have a telephone and the list frame of telephone numbers additionally excluded unlisted telephone numbers. The exclusions of the telephone frame were compensated by the use of the area frame. For example, the telephone list frame was only used for areas where the area frame was also used. Overall, Cycle 1.1 excluded about 2% of the Canadian population.

The frame for Cycle 1.2 was the LFS area frame with the same exclusion as for Cycle 1.1. Cycle 1.2 targeted Canadians 15 years of age and older. Of note for Cycle 1.2, the area frame excluded people who do not have a residence anywhere (*i.e. the homeless*).

Unit Response

Response rates for the CCHS are calculated at the person level. The response rate for Cycle 1.1 was 85.1% for the area frame units, and 83.1% for telephone frame units, for a combined rate of 84.7%. Provincial combined response rates varied from 82% in Ontario to 89.5% in Manitoba, with health region combined response rates varying from 70.2% to 92.3%. The response rate for Cycle 1.2 was 77%, with provincial combined response rates varying from 73.4% in Ontario to 82.4% in Manitoba.

Various strategies were put in place during data collection of Cycles 1.1 and 1.2 to improve response rates such as: supervision and control of the interviewers, targeted interviewer training, use of introductory letters and brochures, use of questionnaires translated in languages other than French and English (*Chinese, Punjabi and Inuktitut*), non-response follow-ups, data quality monitoring and validation, adjustments to collection methods, transfer of caseloads to other interviewers or offices, use of proxy reporting is some circumstances, etc. The systematic data quality monitoring of data collection for Cycle 1.1 for example through the use of data quality management systems consisting of various weekly, monthly, quarterly and semi-annual data quality indicators allowed the detection of several problems that occurred in the field and to put in place measures to rectify the situation such as: the use of the telephone list frame was extended to more health regions halfway through data collection to replace the RDD frame where very low hit rates were observed; higher than anticipated proxy reporting rates were detected and reduced by re-enforcing interviewing procedures; and a problem of unusually short interviews was ascribed to non-adherence of procedures and was corrected *(Béland, Dufour and Hamel 2001*).

The weighting strategy developed for each Cycle of the CCHS involves various adjustments, including household and person level non-response adjustments. The weighting for Cycle 1.1 was first performed individually for each frame, and resulting weights where then combined into a single set of weights (*Brisebois and Thivièrge, 2001*).

Sample Size and Sampling Error

The survey file for Cycle 1.1 includes 131,535 respondents and 36,984 respondents for Cycle 1.2. The file for Cycle 1.1 includes a sample buy-in for Prince-Edward-Island, and the file for Cycle 1.2 includes sample buy-ins for Nova-Scotia and Ontario.

Exact variance and coefficients of variations (CVs) of estimates produced from these data files are calculated using the bootstrap method. Approximate CV look-up tables are also available to users. All CCHS tables on CANSIM and other parts of the Statistics Canada web site include confidence intervals or a standard indicator of the precision of the estimates based on CV's and the Statistics Canada guidelines. For example, estimates with a CV of more than 33.3% are suppressed from the tables and the table cells are marked with an 'F'.

Item Response

Item non-response for both Cycles 1.1 and 1.2 was generally less than 1% and often less than 0.1%. Higher non-response was observed for a few variables for Cycle 1.1 such as weight, PAP test, sexual relations, depression, social support, self-esteem, sedentary activities, work stress, etc., but this was mainly related to higher than expected proxy reporting rates. Some questions were not asked in proxy mode, adding to the item non-

response. Refusal rates were actually very low for all items except income, which has a non-response rate of 10%.

Higher than 1% item non-response was also observed for a few items in Cycle 1.2, but was generally less than 3%. The items that were mostly affected were some of the derived variables for the mental health disorders. No scores were computed if any of the variables necessary for the computation had a non-response. Some of the items asking for the date of an event in the respondent's life also had a higher rate of 'don't know' answer. Some of the other items with higher non-response include weight, alcohol consumption, work stress, and income.

Various actions were taken to keep item non-response to a minimum. Interviewers were trained on the purpose and concepts of all the questions. The information for the CCHS was collected using a computer-assisted approach which ensured that all and only appropriate questions were asked. Item non-response was also monitored at the interviewer level during data collection to ensure adherence to procedures.

Identifying and Correcting Errors

Most editing for the CCHS is conducted at the time of the interview by the Computer Assisted Interview (CAI) application. Some types of inconsistent or unusual reporting were edited after data collection at Head Office. Inconsistencies were usually corrected by setting answers to questions to 'not stated'.

For Cycle 1.1, some questions or modules were appropriate for self-response only and these were skipped for the 6.3 % of interviews completed by proxy. The missing values were imputed using the 'nearest neighbour' method, except for some modules for which data quality could not be improved through imputation *(St-Pierre, M and Béland, Y. 2002*).

Most of the reported information for Cycle 1.2 on mental health was left as is. Because of the potential sensitivity of the content, it was felt inappropriate to question the respondents on inconsistent answers during the interview and the information was left as collected on the file to allow researchers to deal with it as they see fit given that the answers could be subject to different interpretation.

Other Accuracy Issues

Detailed Microdata User Guides were developed for both cycles to provide all the relevant background information on each of the surveys (*background, methodology, data quality, data dictionary, derived variables specifications, etc*). Similar information is and will also be loaded in the Integrated Meta Data Base (IMDB).

Special studies were conducted on the survey data. These include a validation of CCHS results in relation to various other surveys for both CCHS cycles conducted so far. A special study comparing the profile of respondents accepting to share their data with

other departments was conducted. Another study was conducted on the effect of collection mode (*telephone vs face-to-face for Cycle 1.1*) on survey estimates.

As input to the recent audit of health statistics by the Office of the Auditor General (OAG), a self-assessment of data quality assurance was done for the CCHS, followed by a review of the self-assessment. The self-assessment, the review and the OAG concluded that "the survey is well supported by a comprehensive quality assurance program", that "the survey met (*Statistics Canada's*) standards for data quality" and that the survey "maintains data quality through extensive consultations, testing and monitoring, and other data quality assurance activities" (*Report of the Auditor General of Canada to the House of Commons, Chapter 6, Statistics Canada - Managing the Quality of Health Statistics, December 2002*). In addition, it "yielded reliable estimates as a result of the sample size and method at the level of health regions".