Health and Activity Limitation Survey (HALS) 1991

Statistical Data Documentation System
Reference Number 3251

Data Quality
Statements

Health and Activity Limitation Survey - 1991

Data quality

Statistics from the 1991 Health and Activity Limitation Survey (HALS) database are estimates based on a sample survey of a portion of the Canadian population (approximately 1 out of 75 persons in the " yes" sample and 1 out of 200 persons in the "no" sample). As a result, the statistics are subject to two types of errors: sampling and non-sampling errors.

Sampling error is the difference between the estimate derived from a sample and the result that would have been obtained from a population census using the same data collection procedures. For a sample survey such as HALS, this error is estimated from the survey data. The measurement of error used is the standard deviation of the estimate. When a sampling error is more than 33 1/3% of the estimate itself, it is considered to be too unreliable to be published. In such a case, the symbol " -- " appears in statistical tables in place of the estimate. When the sampling error is between 16 2/3% and 33 1/3%, the corresponding estimate is accompanied by the symbol " * ' in a table. Such estimates should be used with caution. Finally, all estimates with a sampling error of less than 16 2/3% can be used without restriction.

All other types of errors (coverage, response, processing and non-response) are called **non-sampling errors.** It is generally difficult to identify and evaluate precisely some of these errors.

Coverage errors arise when there is a difference between the target population and the sample population. Integrating HALS with the census of population has greatly reduced this type of error. Only a certain portion of Indian reserves and collective dwellings were excluded from the sampling process, but because of their small numbers their effect on the total population is negligible. Consequently, coverage errors should not have a significant influence on the HALS data.

A **response error** occurs when the respondent misunderstands a question, and the interviewer records an incorrect answer.

Processing errors may occur at various stages including: coding, data capture and imputation.

All statistical surveys are susceptible to a certain percentage of **non-response** among the selected sample. Total non-response occurs when, for one reason or another, a selected respondent cannot be interviewed. Partial non-response occurs if only part of the questionnaire is completed. The impact of non-response errors on estimates depends on the level of non-response and particularly, on any differences between the characteristics of respondents and non-respondents. In principle, the more marked these differences, the greater the impact on the accuracy of the estimates.

With respect to HALS, the response rate of 87% compares favourably with the rate generally observed for this type of survey. In addition, various methods have been used to reduce the bias caused by non-responses (e.g., adjusting the data to reflect the distribution of certain demographic characteristics obtained from the census).