
II Survey Methodology

Introduction

The Canadian Heart Health Data Base (CHHDB) is an integration of data from ten Provincial Heart Health Surveys conducted between 1986 and 1992. The Provincial Heart Health Surveys were conducted as part of the Canadian Heart Health Initiative and have been a collaborative effort among provincial departments of health, Health Canada and The Heart and Stroke Foundation of Canada. The origin of the collaborative approach to cardiovascular disease prevention (CVD) lies in a report prepared by the Federal Provincial Working Group on Cardiovascular Disease Prevention and Control.¹

The Canadian Heart Health Data Base consists of two sets of integrated data: core information collected by all ten provincial surveys, and family history information collected by only four provinces — Quebec, Ontario Saskatchewan and Alberta.

Objectives

The objective of the surveys was to estimate at the provincial level, the prevalence of CVD risk factors, the knowledge and awareness levels of CVD causes, the consequences of CVD, and the associated risk factors and lifestyle behaviours.

Survey Methodology and Sample Design

All the Provincial Heart Health Surveys were carried out in two stages, using a stratified, multi-stage probability sample design. The sample size for each provincial survey was large enough to provide approximately 2 000 responses to the survey — enough to estimate the prevalence of a risk of 15% or more to within 10% at a 95% reliability level. The number of expected responses by age and sex are given in Table 1. The number of persons who were selected and located, and who responded to two phases of data collection to the provincial surveys are given in Table 2.

¹ Promoting Health in Canada: Report of the Federal Provincial Working Group on Cardiovascular Disease Prevention and Control, Department of National Health and Welfare, Ottawa, 1987.

Table 1: Expected Responses to each Provincial Survey

| AGE/SEX GROUPS | 18-34 | 35-64 | 65-74 | TOTAL |
|----------------|-------|-------|-------|-------|
| MALE | 600 | 300 | 100 | 1000 |
| FEMALE | 600 | 300 | 100 | 1000 |
| TOTAL | 1200 | 600 | 200 | 2000 |

Table 2: Sample Size and Responses

| Province | Number of Persons Sampled | Number of Persons Located | Number of Home Visits | Number of Clinic Visits | Number of Persons who Provided Fasting Blood Sample |
|---------------|---------------------------|---------------------------|-----------------------|-------------------------|---|
| Nfld | 4865 | 3185 | 2394 | 2067 | 2017 |
| PEI | 2711 | 2318 | 2088 | 2026 | 1959 |
| Nova Scotia | 4302 | 2735 | 2108 | 1798 | 1585 |
| New Brunswick | 3867 | 2737 | 2093 | 1948 | 1839 |
| Quebec | 3792 | 3052 | 2353 | 2095 | 2028 |
| Ontario | 5817 | 3639 | 2538 | 2039 | 1765 |
| Manitoba | 4054 | 3597 | 2766 | 2316 | 2164 |
| Saskatchewan | 3895 | 2893 | 2158 | 1749 | 1609 |
| Alberta | 3436 | 2739 | 2237 | 1993 | 1836 |
| BC | 5195 | 2960 | 2394 | 2064 | 1886 |
| TOTAL | 41,934 | 29,855 | 23,129 | 20,095 | 18,688 |

Two Stages of the Surveys

During the first stage of data collection, respondents were visited in their homes by public health nurses to collect basic demographic data, and information about CVD risk factors, attitudes, and opinions on heart health related issues. Two blood pressure readings were taken, one at the beginning and the other at the end of the interview. Before the visit concluded, arrangements were made for the respondents to visit the clinic for the second stage of the data collection. Respondents were asked to fast for at least eight hours, prior to their scheduled clinic visit.

During the clinic visit, anthropometric measurements were taken (including height and weight), along with a blood sample and two blood pressure readings.

Sampling Frame

The Medical Insurance Registers (MIRs) of the provinces were used as the sampling frame or pool from which the sample was drawn. Special procedures were developed to handle the list of selected persons in a way that respected confidentiality requirements.

Use of the MIRs to select the survey sample was advantageous in several ways:

1. The names and addresses of close to 100% of all residents in the provinces were listed.
2. The standard geographic identification was available for each resident. This information facilitated the grouping of residents by geographic areas.
3. The age and sex information for each resident was available, which made it possible to stratify the population by age/sex groups within geographic areas.
4. The cost and time taken to select the required sample were relatively small compared to alternative ways of selecting the sample.

It was not possible to locate all individuals selected from the MIRs as they were not up to date with respect to the deaths and mobility of the population. To minimize bias due to mobility, when the address of the selected person was not current, the person was traced to the new address and if he/she had moved, but was still within the selected area, he/she was included in the survey.

Target Population

The target population included all males and females between the ages of 18-74, living in the province at the time of the survey, but excluded those living on Indian reserves, in military camps, and institutions such as prisons. Mortality due to CVD and the prevalence of CVD risk factors among the population under 18 are marginal and, therefore, were excluded. For operational reasons the population over the age of 74, along with persons on Indian reserves, military camps and in institutions were excluded.

Stratification and Sample Selection

An independent sample was selected for each province using a similar stratified, multi-stage probability sample design. Urban population centres of each province, as defined by the 1986 Canadian Population Census, were stratified into a number of urban strata on the basis of their population size. Rural areas of Standard Geographic areas (e.g. Census Division, Health Units) in each province formed the rural stratum.

From the urban strata of each province, a number of areas were selected with “probability proportional to size” (pps). A number of Standard Geographic areas were also selected with pps from the rural strata to select the sample of persons from rural areas.

For each province, all persons on the MIR within the selected areas were further stratified into six age/sex groups: males/females 18-34, 35-64 and 65-74 years. Independent samples of persons were randomly selected from each age/sex stratum to give the required number of responses as given in Table 1. It is assumed that the surveys would have a response rate of 80% and 80% of the addresses on the MIR were current to calculate the sample sizes. The samples were selected in replicates of equal sizes from each age/sex stratum.

Sample Weights

Notations:

| | | |
|-------------|---|---|
| N_{phai} | — | Number of persons on the MIR of each province “p” in stratum “h”, area “a” and age/sex group “i”; |
| n_{phair} | — | Number of persons selected from province “p”; |
| m_{phair} | — | Number of persons out of (n_{phair}) responded to the home interview; |
| s_{phair} | — | Number of persons out of (m_{phair}) came to the clinic; |

α_{pha} — First stage selection probability factor for area “a” selected from stratum “h” and province “p”;

w_{phair} — Household weight for respondent from (phair);

t_{phair} — Clinic weight for respondent from (phair);

w^*_{phair} and t^*_{phair} are the corresponding “unadjusted” weights;

P_{pi} — Statistics Canada Population estimates (closest to the survey date) of province “p” by age/sex “i”;

\hat{P}_{pi} — Estimate of P_{pi} from the survey;

r — Number of replicates selected from age/sex group “i”;

Calculation of Weights

$$w^*_{phair} = \alpha_{pha} N_{phai} / (m_{phair})$$

and

$$t^*_{phair} = \alpha_{pha} N_{phai} / (s_{phair})$$

$$\hat{P}_{pi} = \sum_{h,a} \alpha_{(pha)} N_{phai}$$

$$w_{phair} = (P_{pi} / \hat{P}_{pi}) w^*_{phair}$$

and

$$t_{phair} = (P_{pi}/\hat{P}_{pi})t^*_{phair}$$

Sampling weights (w_{phair})'s are used for information collected during the home interview. The sampling weights ($t_{(phair)}$)'s are used for information collected during the clinic visit and for analysis of information collected during the home interview and clinic visit, jointly.

Estimation

Household weights were used to tabulate the characteristics collected during the home interview. Clinic weights were used to tabulate the characteristics collected during the clinic visit and also to cross-tabulate the characteristics collected during the home interview and clinic visit.

Tabulation of Categorical Estimates

Estimates of the number of people with a certain characteristic can be obtained from the micro data file by summing the weights of all records possessing the characteristic(s) of interest. Proportions and ratios of the form X/Y are obtained by:

- (a) summing the weights of records having the characteristic of interest for the numerator (X),
- (b) summing the weights of records having the characteristics of interest for the denominator (Y), then,
- (c) dividing the numerator estimate by the denominator estimate.

Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the micro data file by multiplying the value of the variable of interest by the final weight for each record, then summing this quantity over all records of interest. For example, to obtain an estimate of years of education received by people with hypertension, multiply the value of years of education of each person with hypertension by the household weight, then sum the values over all records of persons with hypertension.

To obtain a weighted average of the form X/Y, the numerator (X) is calculated as for a quantitative estimate and the denominator (Y) is calculated as for a categorical estimate. For example, to estimate the *average* years of education received by people with hypertension:

- (a) X, estimate the total years of education for people with hypertension as described above, and
- (b) Y, estimate the number of people with hypertension by summing the household weights of all persons with hypertension.

Then, the average years of education received by persons with hypertension is given by X/Y.

Survey Questionnaires

Personal information and anthropometric measurements with four blood pressure readings and a blood sample were collected by all ten provincial surveys.

Questionnaires used by the ten provinces to collect information during the home visits are provided in the Questionnaire section of this document. The questionnaires were pre-tested before the surveys were conducted.

Edit and Imputation of Missing Data

Quality control procedures were used to achieve high accuracy in coding and capturing data. The capture data then went through an edit procedure to eliminate any inconsistencies in the data set. The complete non-responses by the selected persons to the home interview and/or clinic interviews were accounted for by adjusting the weights of the respondents. Partial non-responses to such items as the demographic questions were imputed, but no imputation was done for attitudinal/opinion/awareness questions, the blood pressure, anthropometric and cholesterol measurements. For blood pressure measurements, the average systolic and diastolic blood pressure were determined on the number of readings available. For quality check of the laboratory procedures, two blood samples for every 50th person was taken. In some cases, if the blood sample of a person was missing, an effort was made to obtain another blood sample. However, the number of partial non-response cases were very few.

Blood Pressure (BP) Measurement

Two diastolic and two systolic blood pressure readings were obtained during the home interview and another two readings during the clinic visits from each respondent. The reading of the blood pressure measurement has a high variability. Generally, only one reading gives an over estimate of blood pressure. Therefore, at least two or more readings, at different time periods, are required to obtain accurate blood pressure measurements. These heart health surveys are one of the first large scale surveys where four readings of blood pressure were obtained. Blood pressure was measured by the nurses using standard mercury sphygmomanometers. An appropriate size cuff was placed on the participant's right arm, with the limb held at the level of the heart. The participant sat quietly for five minutes and did not smoke for 30 minutes prior to measurement. The radial pulse was palpated and counted for fifteen seconds. The first and fifth Korotkoff sounds were recorded for each measurement. For sounds which continued to zero mm Hg the fourth Korotkoff sounds were recorded as the diastolic blood pressure. The nurses were trained in the taking of blood pressure measurements according to standardized procedures of the Lipid Research Clinics Program (National Heart and Lung Institute, 1984). Blood pressure measurements were taken, both at the beginning and at the end of the interview, and at the beginning and end of the clinic visit. The blood pressure data presented on this database is based on the mean of the four measurements.

Height and Weight Measurements

To obtain consistent and accurate measurement of height and weight, a protocol was developed along with a training package to train the nurses to take height and weight measurements. The nurses measured the individual's height and weight using standardized scales. Participants dressed in normal indoor clothing, without shoes. Height measured in centimetres or inches, and weight in kilograms or pounds. The height and weight data converted to Body Mass Index Weight ($\text{kg}/\text{Height (m)}^2$) at the data processing stage.

Analysis of Blood Lipid

The Lipid Research Clinic Laboratory at the University of Toronto was used to carry out the analysis of the blood samples from all the surveys. An 8-hour fasting venous blood sample was obtained using Vacuutainer tubes and tourniquet. Two ten ml lavender-stopper Vacutainer tubes, containing sodium EDTA, were filled as completely as possible. The filled tubes were promptly mixed by inverting the tubes eight times. They were labelled and placed in wet ice pending centrifugation, which was carried out at the clinic. Serum thus obtained was transferred to Bijou bottles by Pasteur pipettes and shipped, on ice, for lipid analysis to the Lipid Research Clinic (LRC) Laboratory at the University of Toronto. The Lipid fractions determined were total cholesterol (TC), triglyceride (TG), high density lipoprotein-cholesterol (HDL), and low density lipoprotein-cholesterol (LDL). The lipid determinations were made according to procedures specified in the LRC Laboratory methods manual, and met strict requirements for standardization.

Data Processing

Ten provincial data bases prepared by respective provinces were integrated to from Canadian Heart Health Data Base (CHHDB) by the Canadian Heart Health Data Base Centre, Division of Community Medicine, Memorial University of Newfoundland. The tabulation specifications were prepared, based on the objectives of the surveys, and were carried out using the appropriate weights.

Data Interpretation Committees

The results of the surveys were examined and interpreted by the Data Interpretation Committee, which consists of internationally-recognized subject matter experts. The survey report, outlining the results and the interpretation of the data of each province, was prepared and published. A discussion of the methodology has been published in a special supplement to the Canadian Medical Association Journal (June 1, 1992).