Microdata User Guide

Maternity Experiences Survey

2006

Statistics Canada Statistique Canada

Canada

Table of Contents

1.0 Introduction 5

2.0 Background 7

3.0 Objectives 9

4.0 Concepts and Definitions 11

5.0 Survey Methodology 13
5.1 Target Population **13**
5.2 Sampling Frame **13**
5.2.1 Population Undercoverage and Overcoverage **13**
5.3 Sample Design **14**
5.3.1 Stratification **14**
5.3.2 Allocation **14**
5.4 Sample Size **14**

6.0 Data Collection 17
6.1 Non-response **18**

7.0 Data Processing 19
7.1 Data Capture **19**
7.2 Editing **19**
7.3 Coding of Open-ended Questions **20**
7.4 Imputation **20**
7.5 Creation of Derived Variables **20**
7.6 Weighting **22**

8.0 Data Quality 23
8.1 Response Rates **23**
8.2 Survey Errors **24**
8.2.1 The Frame **24**
8.2.2 Data Collection **24**
8.2.3 Data Processing **26**
8.2.4 Non-response **28**
8.2.5 Measurement of Sampling Error **29**

9.0 Guidelines for Tabulation, Analysis and Release 31
9.1 Rounding Guidelines **31**
9.2 Sample Weighting Guidelines for Tabulation **31**
9.3 Definitions of Types of Estimates: Categorical and Quantitative **32**
9.3.1 Categorical Estimates **32**
9.3.2 Quantitative Estimates **32**
9.3.3 Tabulation of Categorical Estimates **33**
9.3.4 Tabulation of Quantitative Estimates **33**
9.4 Guidelines for Statistical Analysis **33**
9.5 Coefficient of Variation Release Guidelines **34**
9.6 Release Cut-off's for the Maternity Experiences Survey **36**

10.0 Approximate Sampling Variability Tables 37
10.1 How to Use the Coefficient of Variation Tables for Categorical Estimates **38**
10.1.1 Examples of Using the Coefficient of Variation Tables for Categorical Estimates **39**
10.2 How to Use the Coefficient of Variation Tables to Obtain Confidence Limits **42**
10.2.1 Example of Using the Coefficient of Variation Tables to Obtain Confidence Limits **43**
10.3 How to Use the Coefficient of Variation Tables to Do a T-test **44**
10.3.1 Example of Using the Coefficient of Variation Tables to Do a T-test **44**
10.4 Coefficients of Variation for Quantitative Estimates **45**
10.5 Coefficient of Variation Tables **45**

11.0 Weighting 47

12.0 Questionnaire 51

13.0 Record Layout with Univariate Frequencies 53

1.0 Introduction

The Maternity Experiences Survey (MES) was conducted by Statistics Canada in the fall of 2006 with the cooperation and support of the Public Health Agency of Canada. This manual has been produced to facilitate the manipulation of the microdata file of the survey results.

Any question about the data set or its use should be directed to:

Statistics Canada

Client Services
Special Surveys Division
Telephone: 613-951-3321 or call toll-free 1-800-461-9050
Fax: 613-951-4527
E-mail: ssd@statcan.ca

Public Health Agency of Canada

Maternal and Infant Health Section
Health Surveillance and Epidemiology Division
Jeanne Mance Building, Tenth Floor, 200 Eglantine Driveway, AL 1910D,
Ottawa, Ontario, K1A 0K9
Telephone: 613-941-2395
Fax: 613-941-9927
E-mail: MES@phac-aspc.gc.ca

2.0 Background

The Maternity Experiences Survey (MES) is an initiative of the Canadian Perinatal Surveillance System (CPSS), which is managed by the Maternal and Infant Health Section of the Public Health Agency of Canada. Following its inception, the CPSS identified 52 perinatal health indicators that should be monitored in a comprehensive national perinatal surveillance system. However, existing national data sources such as vital statistics, hospitalization data and current population health surveys only allowed the CPSS to report on 27 of these indicators.

In 1999 the CPSS established a multidisciplinary working group, the Maternity Experiences Study Group, to design a new survey that would provide the missing information. The Study Group assumed the lead role in determining the survey's content and questionnaire design and worked with Statistics Canada on implementation of the first national maternity survey.

The initial work on the development of the survey involving Statistics Canada began in 1999 and resulted in a pilot test carried out in November 2002. The key recommendations for the main survey called for a reduction of content and the use of computer-assisted telephone interviewing application as the collection method. A second pilot of the modified version of the survey was recommended.

The main survey could not take place as originally planned, so consequently the second pilot test was delayed. In the meantime, it was decided to use a different sample frame than initially planned that is the 2006 Census of Population instead of birth registries. The second pilot test took place in the fall of 2005.

The survey was carried out by Statistics Canada on behalf of the Public Health Agency of Canada in the fall of 2006. Close to 6,500 mothers of babies between five and fourteen months of age from the provinces and territories were interviewed over the phone about their pregnancy, birth, and post-partum experiences.

3.0 Objectives

In general terms, the Maternity Experiences Survey (MES) was to address the data gaps by collecting data from recent mothers on important perinatal health indicators, such as:

- recent mothers' views of their care,
- information they had on pregnancy, childbirth and the post-partum period, and
- practices and factors that may affect their maternity experiences.

The survey's definitive goal was to gather information that would be useful for improving the well-being of pregnant women, new mothers and infants in Canada. Of particular interest to the survey were mothers who may have special needs or be at risk with respect to their pregnancy outcome that is:

- teenage mothers – under the age of 20 on the baby's date of birth;
- recent immigrants – landed in Canada at most 5 years before the MES; and
- aboriginal mothers.

The 25 modules of the questionnaire aimed at covering the reproductive history and socio-demographic characteristics of mothers, their behaviour and attitudes before pregnancy, various aspects of their pregnancy, the labour and birth as well as the post-partum period. The specific topics included:

Pregnancy:

- the mother's feelings about pregnancy and timing;
- taking of folic acid;
- smoking, drinking and drug use;
- Body Mass Index (BMI) before pregnancy and weight gain during pregnancy;
- stressful events, physical abuse;
- information about pregnancy and birth;
- access to and use of prenatal care.

Labour and birth:

- institution or home birth;
- healthcare provider who delivered baby;
- presence of spouse and/or companion during labour/birth;
- overall evaluation of birth experience;
- vaginal birth:
  - procedures
  - pain medications
  - medication-free pain control
  - duration of labour
- caesarean birth
  - planned – medical or non-medical reasons
  - unplanned – vaginal birth attempted or not
- baby's gestational age and weight at birth.

Post-partum period:

- mother's contact with the newborn during the first 24 hours after birth;
- use of intensive care, length of hospital stay;
- breastfeeding;
- access to health care for the mother and the infant;
- physical health problems due to birth;
- post-partum depression;
- satisfaction with healthcare providers before, during and after childbirth; and
- return to work.

4.0 Concepts and Definitions

This chapter outlines concepts and definitions of interest to the users. Users are referred to Chapter 12.0 of this document for a copy of the actual survey questionnaire(s) used.

Aboriginal mother - a mother who reported being an Aboriginal person that is North American Indian, Métis or Inuit (SD_Q06).

Amniocentisis - a diagnostic test in which a small amount of amniotic fluid surrounding the baby in the uterus is taken through a needle inserted into the woman's abdomen (SI_Q09).

Birthing ball - this is a large, flexible ball a woman in labour sits on and moves her pelvis/hips on (PM_Q01F).

Doula/paid attendant - a doula is a labour assistant/coach who gives support to women and their families during the pregnancy, birth and post-partum period. A doula is not a healthcare provider (PC_Q05A).

Ectopic pregnancy - also known as a tubal pregnancy. A pregnancy in which the fetus is outside the uterus, most often in the fallopian tubes (i.e. the fertilized egg does not descend into the uterus). An ectopic pregnancy often results in an emergency abdominal operation (RH_Q10).

Epidural - an epidural or spinal anaesthesia involves the injection of medication into the back (into the spinal column via a space between two bones), resulting in numbness of the lower part of the body (BB_Q01 and PM_Q08).

Episiotomy - a cut that a doctor or midwife makes in the perineum or vagina to widen the birth canal and make it easier for the baby to be delivered (BB_Q03).

Healthcare provider - is a trained health professional such as:

- obstetrician
- gynaecologist
- a doctor who is both an obstetrician and gynaecologist (OBGYN)
- family doctor or general practitioner
- midwife
- nurse / nurse practitioner

Note: A doula is NOT a healthcare provider.

Maternity benefits - maternity benefits paid through the employment insurance (EI) program or in Quebec by the provincial government are payable to the birth mother for a maximum of 15 weeks (WA_Q06).

Miscarriage - an expulsion of a fetus from the womb before it can survive independently (generally before the 20th week, or 5th month, of pregnancy). It is also sometimes referred to as a spontaneous abortion. A pregnancy ending in the expulsion of a blighted ovum is to be considered a miscarriage as well (RH_Q02 and RH_Q10).

Parental benefits - parental benefits paid through the employment insurance program or in Quebec by the provincial government are payable to either parent up to a maximum of 35 weeks while they are caring for their baby (WA_Q07).

Perineal tears - spontaneous tears in the tissues of the perineum (vaginal opening) that occur during childbirth (BB_Q03).

Post-partum depression - Feelings of sadness, disinterest, and anxiety that continue for anything from two weeks to many months. The Edinburgh Post-Partum Depression Scale (Cox and Holden, 1987) was

included in the questionnaire (ES_Q01 to ES_Q11).

Premature labour - the onset of labour before the 37th completed week of pregnancy – more than three weeks before the due date (RH_Q07).

Prenatal care - visits to a doctor, nurse or other healthcare provider by a pregnant woman, before the onset of labour and the birth of her baby, to get checkups and advice about the pregnancy. Visits solely for an ultrasound or blood test are not considered to be prenatal care visits (PC_Q01).

Recent immigrant - refer to landed immigrants who came to Canada up to five years prior to a given census year. For the 2006 Census, recent immigrants are landed immigrants who arrived in Canada between January 1, 2001 and Census Day, May 16, 2006.

S.I.D.S. (Sudden Infant Death Syndrome) - is also known as crib death or cot death. It refers to the sudden and unexpected death of an apparently healthy infant (PI_Q03).

Stillbirth - a baby born after the 20th week, or 5th month of pregnancy without any signs of life (RH_Q02 and RH_Q08).

Vacuum extraction (or birth) - also called a suction-assisted birth or suction delivery. A procedure in which a doctor assists the baby in being born by placing a small suction cup on the baby's head to gently pull the baby out (VB_Q01B).

Young mother - a young mother is defined as anyone who was under the age of 20 on the selected baby's date of birth.

5.0 Survey Methodology

This chapter describes the target population, the sampling frame and the sample design of the Maternity Experiences Survey (MES).

5.1 Target Population

The MES target population consisted of biological mothers who were 15 years of age and older at the time of their babies' singleton live birth in Canada, and who also lived with their infants at the time of the survey. In addition, the babies in the target population were required to have been born between February 15, 2006 and May 15, 2006 in the provinces and between November 1, 2005 and February 1, 2006 in the territories. Mothers living on Indian Reserves and in collective dwellings at the time of the 2006 Census were excluded.

5.2 Sampling Frame

The sampling frame for the MES was constructed from the 2006 Canadian Census of Population. The Census is conducted using a short questionnaire (Form 2A) and a long questionnaire (Form 2B/2D). Each household in Canada receives either a Form 2A or a Form 2B/2D questionnaire. In most regions of the country, one in five households is selected at random to receive the Form 2B/2D questionnaire. In the provinces, only the Form 2A sample was used to construct the MES frame. In the territories, both the Form 2A and Form 2B/2D samples were used. Indian Reserves and collective dwellings were excluded.

The MES frame was constructed in several steps: The Census date of birth was used to identify babies born between the target dates. Babies that were obviously part of a multiple birth were dropped. The next step was to identify the mother of each baby on the sampling frame. Because the Census only provides information about a household member's relationship to the first person listed on the questionnaire (Person 1), it was often necessary to examine every female household member's relationship to Person 1 in conjunction with the baby's relationship to Person 1, to deduce which household member was likely to be the mother. When no one in the household was identified as possibly being the mother, the baby was dropped from the frame. The final sampling frame consisted of 58,972 women.

5.2.1 Population Undercoverage and Overcoverage

As a post-censal survey, the MES sampling frame inherited many of the coverage issues of the Census. Individuals missing from the Census caused undercoverage, whereas individuals included twice caused overcoverage. Reporting errors and capture errors in the birth dates (as well as for other information on the frame) could have caused both overcoverage and undercoverage.

Some undercoverage was caused by survey timelines which forced the creation of the MES sampling frame before Census data were fully processed. This undercoverage was corrected through calibration (see Chapter 11.0).

The MES frame overcovered the target population because the Census does not collect all the information needed to determine whether the mother was in-scope for the MES (e.g., birth mother of a baby born in Canada). This overcoverage was eliminated by identifying out-of-scope mothers during collection, then removing them from the sample, and by extension, from the estimated in-scope target population.

A small amount of undercoverage could have resulted from babies who lived with their mothers part-time, but who were not retained on our frame because they were listed in different households at the time of the Census.

5.3 Sample Design

The MES sample had a two phase design. The first phase was the Census sample which was used to create the MES frame. The second phase was a stratified simple random sample selected without replacement from the frame.

5.3.1 Stratification

The primary stratification was defined in terms of the province or territory at the time of the Census and maternal age at the time of giving birth (less than 20, 20 years and over). The strata for mothers aged 20 and over were further sub-stratified:

- in all provinces by maternal age (20 to 29, 30 years and over);
- in all but the three least-populated provinces by whether the mother resided in a census metropolitan area (CMA); and
- also in the three most populous provinces by whether there were other children in the household.

The younger mothers (less than 20) strata were generally too small for further stratification, except in Ontario and Quebec where they were sub-stratified according to whether the mother resided in a CMA.

5.3.2 Allocation

The Kish allocation method adapted for finite populations was used to allocate the sample to the provinces and territories. A parameter of 0.7 was selected, producing a compromise between equal allocation and proportional allocation. The sample was first allocated to the primary strata, and then allocated to the sub-strata proportionally to the population in the sub-strata. Young mothers were oversampled to increase their representation in the sample. The sample was inflated to anticipate non-response – a response rate of 75% was assumed in the provinces and a response rate of 70% was assumed in the territories.

5.4 Sample Size

The following table shows the sample size for each province.

Provinces and Territories	Sample Size
Newfoundland and Labrador	348
Prince Edward Island	225
Nova Scotia	416
New Brunswick	400
Quebec	1,567
Ontario	2,622
Manitoba	465
Saskatchewan	432
Alberta	814
British Columbia	871
Yukon	80

Northwest Territories	136
Nunavut	166
Canada	**8,542**

6.0 Data Collection

Data collection for the Maternity Experiences Survey (MES) was carried out by four regional Statistics Canada offices (Halifax, Sherbrooke, Sturgeon Falls, and Winnipeg). Additionally, the Toronto office was in charge of cases forwarded to them where interviews had to be conducted in selected languages other than English or French.

Altogether some 60 interviewers, all of them women with substantial interviewing experience, and 10 managers/supervisors worked on the survey. They received about two days of classroom training to become familiar with the survey concepts and the computer application.

In the provinces, data was collected using a computer-assisted telephone interviewing (CATI) application, while in the territories if a telephone interview was not feasible, a personal interview with a paper version of the questionnaire was used where possible. Most of the interviews in the territories were conducted over the telephone. About 30 cases were identified as requiring a personal interview and, consequently, a paper version of the questionnaire. The questions as well as the data sharing agreement were translated into three Inuit dialects.

Before starting the interviews, the regional offices mailed an introductory letter to the sampled mothers. The letter introduced the survey and asked for their cooperation. It included a toll-free number to call for more information as well as the address of Statistics Canada's website where information for MES participants could be found. A pamphlet entitled: *What Mothers Say: the Maternity Experiences Survey* produced by the survey's sponsor – the Public Health Agency of Canada – was enclosed. It contained a brief description of the survey, stated its importance and encouraged mothers to take part. In the provinces, a bilingual English/French pamphlet was distributed, while in the North a special version in English and Baffin Syllabic with pictures of Inuit mothers and babies was enclosed.

Data collection for the MES began in the provinces as planned, on October 23, 2006. It ended as planned on the 8th of December in the Halifax and Sherbrooke offices, but was extended by two more days i.e. until December 10th in Sturgeon Falls and Winnipeg to obtain some more interviews, particularly with young mothers in Ontario and in British Columbia.

Data collection in the territories was delayed until the beginning of November because the Census frame information was not available and the territorial samples could not be selected on time. The collection was extended there until the end of January to benefit from personal interviewing conducted by the other post-censal surveys which were still in the field.

Recent immigrants, that is, those who came to Canada in the past five years, were one of the groups of special interest to the survey. To obtain an accurate picture of their experiences it was important to minimize the number of non-interviews due to language problems. The cases coded by the regional offices as "language barrier" were then completed by the Toronto interviewers when possible. The Toronto regional office employs several interviewers fluent in languages spoken by many recent immigrants. They also received the MES training and were asked to conduct interviews with immigrant women who needed full or partial interpretation. These interviewers had to translate the questions and answers as needed. A glossary of MES terms was available in the languages deemed to be the most in demand, based on the Census and other survey information. Interviewers at the Toronto office completed 150 interviews in 13 different languages. Still, there were cases where attempts to reach mothers who were identified as needing interpretation were not successful (about 30 cases). Additionally, due to a planning mistake some 50 cases that should have been transferred to Toronto were sent there too late to proceed with the interviews.

The telephone interviews lasted, on average, 45 minutes.

6.1 Non-response

The computer application had special features to maximize the success rate of attempts to reach respondents. The 25 allowed calls per case were distributed between working days and the weekend and between mornings, afternoons and evenings taking into consideration the schedules of mothers of young babies.

An effort was made to reach members of target groups of special interest to data users. Because the sample file contained information about the age of the person assumed to be the mother of the selected baby, it was possible to identify young mothers, that is, women under the age of 20 on the baby's date of birth. Membership in another target group, recent immigrants, although not known from the sample file could be indirectly targeted by a combination of information about the first language learned (other than English or French) and residence in Vancouver or Toronto where the percentage of recent immigrants is the highest. Sampled mothers identified this way could, if needed (depending on the success rates in reaching the target groups during the collection) be assigned a priority for contact.

Non-response to the Maternity Experiences Survey resulted mainly from an inability to establish contact with some mothers. Although efforts were made to trace mothers who moved or for whom the Census information was incorrect, at the end of data collection a substantial number of cases were still unresolved.

Because eligibility for participation in the survey had to be confirmed during the screening part of the interview, and that lack of contact did not permit screening (or full screening), the total number of non-response cases (as opposed to out-of-scope cases for the survey) had to be estimated.

Despite the length of the interview refusals were rare. Interviewers were instructed to make all reasonable attempts to obtain the MES interviews by offering to call back at a more convenient time and by explaining the importance of their participation. This was followed by a second call from a senior interviewer.

7.0 Data Processing

The main output of the Maternity Experiences Survey (MES) is a “clean” microdata file. This chapter presents a brief summary of the processing steps involved in producing this file.

7.1 Data Capture

Responses to survey questions are captured directly by the interviewer at the time of the interview using a computerized questionnaire. The computerized questionnaire reduces processing time and costs associated with data entry, transcription errors and data transmission. The response data are encrypted to ensure confidentiality and sent via modem to the appropriate Statistics Canada Regional Office. From there they are transmitted over a secure line to Ottawa for further processing.

Some editing is done directly at the time of the interview. Where the information entered is out of range (too large or small) of expected values, or inconsistent with the previous entries, the interviewer is prompted, through message screens on the computer, to modify the information. However, for some questions interviewers have the option of bypassing the edits, and of skipping questions if the respondent does not know the answer or refuses to answer. Therefore, the response data are subjected to further edit and imputation processes once they arrive in head office.

About 30 of the MES interviews in the territories were conducted in person with the use of a paper questionnaire. The completed questionnaires were sent to the Winnipeg Regional Office and the answers were entered using the CADE (computer-assisted data entry) version of the MES computerized questionnaire. This version is exactly the same as the one used for the computer-assisted telephone interviews except that there are no edits. However, all the MES data, regardless of how entered, were subjected to further edits at the data processing stage.

7.2 Editing

The first stage of survey processing undertaken at head office was the replacement of any “out-of-range” values on the data file with blanks. This process was designed to make further editing easier.

The first type of error treated was errors in questionnaire flow, where questions which did not apply to the respondent (and should therefore not have been answered) were found to contain answers. In this case a computer edit automatically eliminated superfluous data by following the flow of the questionnaire implied by answers to previous, and in some cases, subsequent questions.

The second type of error treated involved a lack of information in questions which should have been answered. For this type of error, a non-response or “not-stated” code was assigned to the item.

The following standard Statistics Canada codes are used on the file:

6, 96, 996, etc. = Valid skip
7, 97, 997, etc. = Don't know
8, 98, 998, etc. = Refused
9, 99, 999, etc. = Not stated

7.3 Coding of Open-ended Questions

A few data items on the questionnaire were recorded by interviewers in an open-ended format. A total of 10 partially or completely open-ended questions were included in the survey. The two completely open-ended questions in the MES asked for the name of the hospital or birthing centre and its location. A list of names of institutions provided by the Public Health Agency of Canada was used for coding.

Eight questions included an "Other - Specify" category (PC_Q03, SI_Q10, BH_Q03, MH_Q05, PI_Q11, SD_Q10, WA_Q01 and WA_Q08). Altogether about 400 entries that fit into existing response categories were recoded during processing. For example, in case of question PC_Q03 which asked "*What prevented you from getting prenatal care as early as you wanted?"* the response recorded as "Doctor too busy to take (me) earlier" was recoded to the first answer category (not read to respondent) "Doctor/healthcare provider unavailable". On the other hand, the answer "Did not want parents to know" remained coded as "Other". No new response categories were created.

7.4 Imputation

Imputation is the process that supplies valid values for those variables that have been identified for a change either because of invalid information or because of missing information. The new values are supplied in such a way as to preserve the underlying structure of the data and to ensure that the resulting records will pass all required edits. In other words, the objective is not to reproduce the true microdata values, but rather to establish internally consistent data records that yield good aggregate estimates.

There was no imputation done for the 2006 Maternity Experiences Survey.

7.5 Creation of Derived Variables

A number of data items on the microdata file have been derived by combining items on the questionnaire in order to facilitate data analysis. Several of them are conversions either from Metric to Imperial units (and vice versa) or from responses given in various time units (days or weeks or months) to one type of unit such as weeks. For example, BHD05 and MHD23 converts the baby's age at the time of the baby's or mother's hospitalization to days when it was reported in weeks or months using the formula i.e., days = months * 4.33 * 7.

Other types of derived variables combine the responses from two or more questions asked of respondents who followed different paths of questioning, such as in case of SDDEDUC (mother's highest level of education completed) which is based on detailed education questions SD_Q11 to SD_Q14.

Another example of a derived variable combining information from two or more questions is gestational age – GEST_AGE (in weeks) which is calculated based on due date information (which assumes a gestational age of 40 weeks) and the reported birth date of the baby. For this variable rounding down is used.

Body Mass Index (BMI) is a comparison of weight relative to the height of respondents. BMI is a derived variable calculated by dividing weight in kilograms by height in metres squared.

An index type of derived variable was created for a measure of the number of kinds of stressful events experienced by respondents in the 12 months before giving birth. The variable SEDSTRSS was based on answers to a set of "Yes" / "No" questions SE_Q02 to SE_Q14. The total of "Yes" responses was calculated for all records that had valid answers to the entire set of questions.

The Edinburgh Post-Partum Depression scale was included in the MES questionnaire. Each of the 10 items had four response categories to choose from. They were scored 0, 1, 2, or 3, with 3 indicating the highest intensity of potentially depressed moods. A total score of 13 or above is generally regarded as an indicator of clinical depression

Census Metropolitan Area (CMA) or Census Agglomeration (CA)
The derived variable CMACA is based on postal code information recoded according to the 2001 Standard Geographical Classification (SGC) - Statistics Canada's official classification for geographical areas in Canada.

The general concept of these standard units is one of an urban core, and the adjacent urban and rural areas that have a high degree of social and economic integration with that urban core, as measured by commuting flows derived from Census of Population data on place of work.

According to the 2001 definition to form a census metropolitan area, the urban core must have a population of at least 100,000. To form a census agglomeration, the urban core must have a population of at least 10,000. Several changes have occurred since 2001:

There are 33 CMAs and 111 CAs in 2006. Six CAs from the previous census became CMAs:

- Moncton in New Brunswick;
- Barrie, Brantford, Guelph, and Peterborough in Ontario; and
- Kelowna in British Columbia.

Seven new CAs were created:

- Bay Roberts in Newfoundland and Labrador;
- Miramichi in New Brunswick;
- Centre Wellington and Ingersoll in Ontario;
- Okotoks and Canmore in Alberta; and
- Salmon Arm in British Columbia.

The 2001 CA of Magog merged with the CMA of Sherbrooke since the commuting interchange between the CMA and CA is equal to at least 35% of the employed labour force living in the CA, based on the 2001 place of work data. Finally, two CAs, Gander and Labrador City in Newfoundland and Labrador, were retired because the population of their urban core dropped below 10,000 in 2001.

While census metropolitan areas and census agglomerations contain approximately 80% of the population of Canada, they cover only 4% of the land area. The census **m**etropolitan area and census agglomeration **i**nfluenced **z**one (MIZ) is a concept that geographically differentiates the area of Canada outside CMAs and CAs. Census subdivisions outside CMAs and CAs are assigned to one of four categories according to the degree of influence (strong, moderate, weak or no influence) that the CMAs and CAs collectively have on each of them. Census subdivisions with the same degree of influence, that is based on the percentage of their resident employed labour force that have a place of work in the urban core(s) of CMAs and CAs, tend to be clustered into zones around the CMAs and CAs.

With the introduction of the MIZ concept, it is possible to classify all census subdivisions in the new Statistical Area Classification (SAC) for data dissemination purposes. Indeed, the SAC classifies census subdivisions according to whether they are a component of a census metropolitan area, a census agglomeration, a census metropolitan area and census agglomeration influenced zone (strong MIZ, moderate MIZ, weak MIZ or no MIZ), or the territories (Nunavut, Northwest Territories and Yukon Territory). The application of this classification to census subdivision data could reveal previously hidden details and help users to study the diversity of non-CMA/CA areas of Canada.

7.6 Weighting

The principle behind estimation in a probability sample is that each person in the sample "represents", besides himself or herself, several other persons not in the sample. For example, in a simple random 2% sample of the population, each person in the sample represents 50 persons in the population.

The weighting phase is a step which calculates, for each record, what this number is. This weight appears on the microdata file, and **must** be used to derive meaningful estimates from the survey. For example if the number of mothers under the age of 20 on the baby's date of birth is to be estimated, it is done by selecting the records referring to those individuals in the sample with that characteristic and summing the weights entered on those records.

Details of the method used to calculate these weights are presented in Chapter 11.0.

8.0 Data Quality

8.1 Response Rates

The following table summarizes the response rates to the Maternity Experiences Survey (MES).

Provinces and Territories	Total Sample	Resolved	Estimated In-scope	Respondents	Resolved Rate (%)	Response Rate (%)
Newfoundland and Labrador	348	304	335	279	87.4	83.3
Prince Edward Island	225	193	220	184	85.8	83.6
Nova Scotia	416	368	398	344	88.5	86.4
New Brunswick	400	335	388	303	83.8	78.1
Quebec	1,567	1,314	1,534	1,256	83.9	81.9
Ontario	2,622	2,059	2,534	1,858	78.5	73.3
Manitoba	465	370	448	341	79.6	76.1
Saskatchewan	432	353	421	341	81.7	81.0
Alberta	814	679	799	651	83.4	81.5
British Columbia	871	682	841	631	78.3	75.0
Yukon	80	62	79	61	77.5	77.2
Northwest Territories	136	107	117	89	78.7	76.1
Nunavut	166	111	130	83	66.9	63.8
Canada	**8,542**	**6,937**	**8,244**	**6,421**	**81.2**	**77.9**

The columns in the table above are defined as follows:

Total Sample
The total number of records that were selected from the sampling frame.

Resolved
The number of cases in which the screening process for determining eligibility was completed during data collection. These cases were confirmed as either in-scope or out-of-scope.

Estimated In-Scope
An estimate of the total number of in-scope cases. It was estimated by calculating the proportion of in-scope cases among the resolved cases and assuming the same proportion of in-scope cases among the unresolved cases. This estimation was carried out within weighting classes to account for different proportions in different classes.

Respondents
The number of cases with sufficient useable data and the respondent's permission for Statistics Canada to share their data with the Public Health Agency of Canada.

$$\text{Resolved rate} = \frac{\text{Resolved}}{\text{Total sample}}$$

$$\text{Response rate} = \frac{\text{Respondents}}{\text{Estimated In-scope}}$$

8.2 Survey Errors

The estimates derived from this survey are based on a sample of households. Somewhat different estimates might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used in the survey. The difference between the estimates obtained from the sample and those resulting from a complete count taken under similar conditions, is called the sampling error of the estimate. Errors which are not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort were taken to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures include the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimized, and coding and edit quality checks to verify the processing logic.

8.2.1 The Frame

The MES frame inherited many of the coverage issues of the Census. Statistics Canada conducts postcensal coverage studies after each Census. Preliminary results for the 2006 Census will be available in March 2008, and the final results in September 2008. In 2001, Census undercoverage was estimated at 3.1%. Data users should be aware that individuals undercovered by the Census may have characteristics that differ from the general population. Frame coverage is also discussed in Section 5.2.1.

8.2.2 Data Collection

Interviewer training consisted of reading the MES Interviewer's Manual, attending about a two day long classroom training session and practicing with the MES training cases on the computer. A description of the background and objectives of the survey was provided, as well as a glossary of terms and a set of questions and answers.

The following are the main clarifications and instructions provided during the classroom training session that went beyond what was included in the Interviewer's Manual:

- In questions CB_Q04 *(In the 3 months before you got pregnant with ^baby's name, did you take a multivitamin containing folic acid or a folic acid supplement?)* and CB_Q06 *(During the first 3 months of your pregnancy with ^baby's name, did you take a multivitamin containing folic acid or a folic acid supplement?)* a "Yes" answer implies that the woman knew (not just assumed) that she was taking folic acid.

- The question MH_Q22 *(Not counting the labour and the birth, have you stayed in a hospital overnight since ^baby's name was born?)* refers to hospitalization due to the mother's health problems.

- The question BB_Q05 *(Did you experience any complications or health problems during labour or the birth that required you to have special care, extra tests, or stay in a hospital?)* should be answered "Yes" if the respondent had an

unexpected caesarean section and “No” if it had been planned and went as expected. Interviewers were reminded that women who had a C-section but attempted a vaginal birth were also posed questions referring to those attempts.

- If the baby’s weight was reported in grams, the minimum value set by the application was 1,000. When the baby’s weight was even lower, this minimum value should be entered with a note stating the actual weight.

- Similarly, the distance to the hospital, had a maximum of 995 (miles or kilometers) which in some instances would not be sufficient. Interviewers were instructed to enter “995” and to attach a note with the actual distance.

- The question on maternity/paternity benefits had to be modified by interviewers conducting the MES in Quebec as these benefits were no longer paid by the Employment Insurance program but by the provincial government.

When the MES interviews began, a need arose for additional instructions on how to deal with certain limitations of the survey’s computer-assisted telephone interviewing application and how to interpret a few of the survey questions. Also, the observers - the MES team members who visited all the regional offices and listened to several interviews - made suggestions that had to be communicated to all interviewers. The Survey Operations staff at head office, were in charge of sending these instructions to all the regional offices.

During data collection the regional offices made the following inquiries:

- How to code in question PC_Q07B *(Did you attend these classes in...?)* the response that prenatal or childbirth education classes taken at CLSCs (an acronym for health clinics in the province of Quebec) - (A health clinic);

- How to interpret a response to question PT_Q05 *(How involved were you in deciding whether or not to have a test for HIV?)* that a test for HIV was a routine test - (probe whether respondent felt she was involved, that she had a say);

- In question HP_Q01 (*Before your pregnancy, did you have any medical conditions or health problems that required you to take medication for more than 2 weeks, have special care, or extra tests during your pregnancy?)* should infertility treatments be considered ”medical conditions” - (No);

- In question PC_Q04 *(How many prenatal care visits did you have?)* should visits solely for an ultrasound or blood tests be considered ”prenatal care visits” - (No);

- The question about identity SD_Q08 *(What is your ethnic or cultural identity?)* was it about identity based on ancestry or any kind of perception of herself reported by the respondent. As the question asked how the respondent perceived herself and to what group she felt she belonged, this could be in terms of ancestry or not.

The collection period for most of the provinces ran from the week of October 23 to December 8th, 2006.

8.2.3 Data Processing

During processing some records were discarded either because respondents did not complete the core of the questionnaire, namely those modules covering conception through to the post-partum period or they did not give Statistics Canada permission to share their answers with the Public Health Agency of Canada. There were 282 such records and these were coded as non-response.

Data processing of the MES was done in a number of steps including clean-up, pre-editing, coding, editing, estimation, and weighting. At each step a picture of the output files is taken and an easy verification can be made comparing files at the current and previous step. This greatly improved the data processing stage.

During data processing some inconsistencies in the data were observed and certain reported values were judged impossible. Action was taken for the following three variables:

PC_Q08Y What was the expected or due date for the birth of ^baby's name? Year

Three cases with an incorrect year were identified and corrected.

GEST_AGE Gestational age of selected baby in weeks

In 158 cases the value for GEST_AGE and due date i.e., PC_Q08D (day), PC_Q08M (month) and PC_Q08Y (year) were set to “Not stated”. These cases were identified by combining gestational age and birth weight and comparing them to the new Canadian standard (Kramer et al) for 22 to 43 weeks. For each gestational age, cases where weight varied from the mean birth weight by more than 2 * IQR were considered impossible. The interquartile range (IQR) is the difference between the upper quartile (Q_3) and the lower quartile (Q_1).

A further 12 records with gestational age greater than 45 weeks had the gestational age and due date variables set to ”Not stated”.

WAD03 Number of weeks pregnant when respondent left work

In 13 cases where the gestational age reported was equal to 43 weeks or more, WAD03 and related variables were set to ”Not stated”.

In most cases where inconsistencies between variables were identified no action was taken because it was not clear which of the variables involved contained erroneous data. An account of these data quality issues is provided below:

Introduction of liquids and solids

BFD08 Age of baby when other liquids were added to feeds
BFD09 Age of baby in days when solids were added to feeds
BFD11 Age of baby in days when last breastfed

There were 349 cases where the respondent reported she stopped breastfeeding her baby before she added other liquids to her baby’s diet. There were also cases where the introduction of solids precedes introduction of liquids. These inconsistencies could be due to a number of things: respondent recall error, respondent misinterpretation of the question, or an entry error by the interviewer.

Mother's hospital length of stay and readmission age

PP_Q12A How many days, weeks or months did you stay in the hospital or clinic after ^baby's name was born?

MH_Q23A How old was ^baby's name the first time you required overnight hospitalization?

There are four cases where the mother's length of stay at the hospital after birth was higher than the age of the baby when the mother was readmitted.

Baby's hospital length of stay and readmission age

PPD15 Length of the baby's stay at the hospital or birthing centre in days

BHD05 Age of baby in days when first hospitalized

There are 23 cases where the newborn's length of stay after birth was higher than the age of the baby at readmission. In 18 of these cases babies were in intensive care or a special care unit (Neonatal Intensive Care Unit) and this might have caused some confusion for the mother answering these questions.

Hospital length of stay and timing of home contact

PPD12 Length of the respondent's stay at the hospital or birthing centre in days

MH_Q02 How old, in days, was ^baby's name when a healthcare provider first contacted you at home?

There are 169 cases where the mother's length of stay after birth was higher than the age of the baby at the time of home contact.

Weight gain during pregnancy

HWD03IMP Weight gained by mother during pregnancy in pounds

There were 30 cases with weight gain in excess of 100 pounds. In some of these cases, weight gain values were similar to pre-pregnancy and post-partum weights, suggesting that women might have provided their actual weight during pregnancy rather than their weight gain.

Gestational age of baby at the time respondent stopped working

There are over 100 cases where the respondent stopped working at a very early gestational age, e.g. less than six weeks and 55 cases where the gestational age when the mother stopped working is higher than the gestational age at birth.

Preterm births

PREMATUR The number of premature births a respondent has had

There are 48 cases where the mother reported no preterm births i.e., PREMATUR = 00, but the gestational age of the baby, GEST_AGE, calculated using the due date and the date of birth of the baby, was less than 37 weeks. There were 22 cases where the gestational age was 37 weeks or more, yet the mother reported all her births were preterm.

Number of pregnancies

BIRTH	How many times have you given birth to a live baby?
STILLBTH	How many pregnancies ended in the birth of a stillborn baby?
ECTOPIC	How many pregnancies ended in a tubal or ectopic pregnancy?
MISCARRY	How many pregnancies ended in a miscarriage?
ABORTION	How many pregnancies ended in a therapeutic or induced abortion?

There are 21 cases in which the respondent's pregnancies are not all accounted for by reported live births, stillbirths, miscarriages, ectopic pregnancies and therapeutic or induced abortions. Some of the discrepancies may be explained by a previous multiple birth – information not captured by the MES questionnaire.

Age at first birth

RH_Q05	How old were you when you gave birth to a live baby for the first time?
AGEATBTH	Age of respondent at birth of selected baby

The reported incidence of same age at first birth and the MES birth (36 cases) seems to be much higher than expected based on estimates derived from hospital birth data. The reason for the higher incidence in the MES file is unclear.

8.2.4 Non-response

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response. Total non-response occurred because the interviewer was either unable to contact the respondent, or the respondent refused to participate in the survey. Total non-response was handled by adjusting the weight of individuals who responded to the survey to compensate for those who did not respond.

In most cases, partial non-response to the survey occurred when the respondent did not understand or misinterpreted a question, refused to answer a question, or could not recall the requested information.

The characteristics of the 1,887 unresolved and non-respondent cases were compared to the characteristics of the 6,421 respondents using the Census data available on the MES frame. For this evaluation, the Census data were weighted using the survey design weights. Differences between the two sets of estimates are an indication of the potential differences in the characteristics of the two groups. A third set of estimates was created from the respondents by combining their Census data with the final MES weights. These final weights incorporate adjustments to the design weights to account for the unresolved cases and the non-respondents (see Chapter 11.0), with the goal of reducing the potential non-response bias. The extent to which this reduction in potential bias was successful is judged by comparing this third set of estimates to a final set of estimates based on the entire MES frame. The results are given in the table below.

Variables		Distribution (percentage) for each Source of Estimates			
		Unresolved / Non-respondents	Respondents (Design weights)	Respondents (Final weights)	Entire MES Frame
Mother's age at birth	Less than 20	5.3	2.6	3.2	3.3
	20 to 29	47.0	44.4	45.3	44.6
	30 to 39	41.7	49.4	48.0	47.4
	40 and over	7.9	3.6	3.6	4.7
Mother's first language	English	58.3	61.4	61.6	61.0
	French	14.7	23.6	21.5	21.7
	Other	27.0	15.0	16.9	17.2
Place of residence	Non-CMA	27.4	33.7	32.5	32.2
	CMA	47.3	50.9	50.1	50.0
	Toronto	25.3	15.4	17.5	17.8
Mother's marital status	Never married	34.8	30.4	29.6	30.8
	Other	65.2	69.6	70.4	69.2
Number of adults in household (aged 18 and over)	Less than 2	11.8	4.2	4.2	5.7
	2 and over	88.2	95.8	95.8	94.3
Other children (aged 0 to 15)	No	40.1	42.9	42.8	42.3
	Yes	59.9	57.1	57.2	57.7
Baby's gender	Male	52.4	51.8	51.7	51.6
	Female	47.6	48.2	48.3	48.4

The results of this study show that the characteristics of the non-respondents and unresolved cases tend to be different from those of the respondents. The study also shows that the respondents, after the weighting adjustments for non-response, appear to be fairly representative for the variables available for the study.

To illustrate this conclusion, consider the proportion of cases where the Mother's first language is neither English nor French: The proportion is 27.0% for the non-respondent and unresolved cases, and 15.0% for the respondents using the design weights. After the weighting adjustments, the proportion is 16.9% for the respondents compared to 17.2% for the MES frame. This shows that although the proportion was quite different for the non-respondent / unresolved cases compared to the respondent cases, the weighting adjustments helped compensate for this difference, and improved the representation of the sample.

8.2.5 Measurement of Sampling Error

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. This section of the documentation outlines the measures of sampling error which Statistics Canada commonly uses and which it urges users producing estimates from this microdata file to use also.

The basis for measuring the potential size of sampling errors is the standard error of the estimates derived from survey results.

However, because of the large variety of estimates that can be produced from a survey, the standard error of an estimate is usually expressed relative to the estimate to which it pertains. This resulting measure, known as the coefficient of variation (CV) of an estimate, is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate.

For example, suppose that, based upon the survey results, one estimates that 26.3% of births in Canada are by caesarean section, and this estimate is found to have a standard error of 0.0058. Then the coefficient of variation of the estimate is calculated as:

$$\left(\frac{0.0058}{0.263} \right) X \ 100 \ \% \ = \ 2.2\%$$

There is more information on the calculation of coefficients of variation in Chapter 10.0.

9.0 Guidelines for Tabulation, Analysis and Release

This chapter of the documentation outlines the guidelines to be adhered to by users tabulating, analyzing, publishing or otherwise releasing any data derived from the survey microdata files. With the aid of these guidelines, users of microdata should be able to produce the same figures as those produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

9.1 Rounding Guidelines

In order that estimates for publication or other release derived from these microdata files correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates:

a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.

b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.

c) Averages, proportions, rates and percentages are to be computed from unrounded components (i.e. numerators and/or denominators) and then are to be rounded themselves to one decimal using normal rounding. In normal rounding to a single digit, if the final or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is increased by 1.

d) Sums and differences of aggregates (or ratios) are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding.

e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released which differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s).

f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

9.2 Sample Weighting Guidelines for Tabulation

The sample design used for the Maternity Experiences Survey (MES) was not self-weighting. When producing simple estimates including the production of ordinary statistical tables, users must apply the proper survey weights.

If proper weights are not used, the estimates derived from the microdata files cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada.

Users should also note that some software packages may not allow the generation of estimates that exactly match those available from Statistics Canada, because of their treatment of the weight field.

9.3 Definitions of Types of Estimates: Categorical and Quantitative

Before discussing how the MES data can be tabulated and analyzed, it is useful to describe the two main types of point estimates of population characteristics which can be generated from the microdata file for the MES.

9.3.1 Categorical Estimates

Categorical estimates are estimates of the number, or percentage of the surveyed population possessing certain characteristics or falling into some defined category. The number of vaginal births is an example of such an estimate. An estimate of the number of persons possessing a certain characteristic may also be referred to as an estimate of an aggregate.

Examples of Categorical Questions:

Q: Did you have a vaginal or caesarean birth for ^baby's name?
R: Vaginal / Caesarean

Q: Immediately after birth, was ^baby's name admitted to an intensive care or special care unit?
R: Yes / No

9.3.2 Quantitative Estimates

Quantitative estimates are estimates of totals or of means, medians and other measures of central tendency of quantities based upon some or all of the members of the surveyed population. They also specifically involve estimates of the form $\hat{X} / \hat{Y}$ where $\hat{X}$ is an estimate of surveyed population quantity total and $\hat{Y}$ is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate is the average number of prenatal care visits. The numerator is an estimate of the total number of prenatal care visits, and its denominator is an estimate of the number of mothers who had prenatal care visits.

Examples of Quantitative Questions:

Q: How many prenatal care visits did you have?
R: |_|_| visits

Q: How many years of formal education have you completed starting with grade one and not counting repeated years at the same level?
R: |_|_| years

9.3.3 Tabulation of Categorical Estimates

Estimates of the number of people with a certain characteristic can be obtained from the microdata file by summing the final weights of all records possessing the characteristic(s) of interest. Proportions and ratios of the form $\hat{X}/\hat{Y}$ are obtained by:

a) summing the final weights of records having the characteristic of interest for the numerator ($\hat{X}$),
b) summing the final weights of records having the characteristic of interest for the denominator ($\hat{Y}$), then
c) dividing estimate a) by estimate b) ($\hat{X}/\hat{Y}$).

9.3.4 Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the microdata 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 the total number of prenatal care visits in Ontario multiply the value reported in question PC_Q04 (number of prenatal care visits) by the final weight for the record, then sum this value over all records with PC_Q06 = 35 (received most of prenatal care in Ontario).

To obtain a weighted average of the form $\hat{X}/\hat{Y}$, the numerator ($\hat{X}$) is calculated as for a quantitative estimate and the denominator ($\hat{Y}$) is calculated as for a categorical estimate. For example, to estimate the average number of prenatal care visits in Ontario,

a) estimate the total number of prenatal care visits in Ontario ($\hat{X}$) as described above,
b) estimate the number of mothers ($\hat{Y}$) in this category by summing the final weights of all records with PC_Q06 = 35, then
c) divide estimate a) by estimate b) ($\hat{X}/\hat{Y}$).

9.4 Guidelines for Statistical Analysis

The MES is based upon a complex sample design, with stratification, multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used. In order for survey estimates and analyses to be free from bias, the survey weights must be used.

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures may differ from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variances that are calculated are poor. Approximate variances for simple estimates such as totals, proportions and ratios (for qualitative variables) can be derived using the accompanying Approximate Sampling Variability Tables.

For other analysis techniques (for example linear regression, logistic regression and analysis of variance), a method exists which can make the variances calculated by the standard packages more meaningful, by incorporating the unequal probabilities of selection. The method rescales the weights so that there is an average weight of 1.

For example, suppose that analysis of all mothers under the age of 20 on the baby's date of birth is required. The steps to rescale the weights are as follows:

1) select all records from the file where the mother was under the age of 20 on the baby's date of birth (AGEGRBTH = 01);

2) calculate the AVERAGE weight for these records by summing the original weights from the microdata file for these records and then dividing by the number of records where the mother was under the age of 20 on the baby's date of birth;

3) for each of these respondents, calculate a RESCALED weight equal to the original weight divided by the AVERAGE weight;

4) perform the analysis for these respondents using the RESCALED weight.

However, because the stratification of the sample's design is still not taken into account, the variance estimates calculated in this way are likely to be incorrect.

It is recommended that the bootstrap method be used to calculate variance estimates for the MES. The bootstrap method is a resampling method for calculating valid variance estimates for complex sampling designs and complex estimators. Bootstrap weights have been produced for the MES and are provided with the data file. Statistics Canada has developed a program called Bootvar which computes variance estimates using bootstrap weights for totals, ratios, differences between ratios, and linear and logistic regressions. The program and documentation can be downloaded from the Statistics Canada website at "www.statcan.ca\English\rdc\whatdata.htm#tools".

9.5 Coefficient of Variation Release Guidelines

Before releasing and/or publishing any estimates from the MES, users should first determine the quality level of the estimate. The quality levels are *acceptable, marginal* and *unacceptable.* Data quality is affected by both sampling and non-sampling errors as discussed in Chapter 8.0. However for this purpose, the quality level of an estimate will be determined only on the basis of sampling error as reflected by the coefficient of variation (CV) as shown in the table below. Nonetheless users should be sure to read Chapter 8.0 to be more fully aware of the quality characteristics of these data.

First, the number of respondents who contribute to the calculation of the estimate should be determined. If this number is less than 30, the weighted estimate should be considered to be of unacceptable quality.

For weighted estimates based on sample sizes of 30 or more, users should determine the coefficient of variation of the estimate and follow the guidelines below. These quality level guidelines should be applied to rounded weighted estimates.

All estimates can be considered releasable. However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

Note that the requirement of a sample size of 30 can be ignored when the coefficients of variation (CV) are estimated using more exact variance estimation techniques such as replication methods, as opposed to approximate variances obtained from the CV tables discussed in Chapter 10.0.

Quality Level Guidelines

Quality Level of Estimate	Guidelines
1) Acceptable	Estimates have a sample size of 30 or more (if CV tables are used), and low coefficients of variation in the range of 0.0% to 16.5%. No warning is required.
2) Marginal	Estimates have a sample size of 30 or more (if CV tables are used), and high coefficients of variation in the range of 16.6% to 33.3%. Estimates should be flagged with the letter E (or some similar identifier). They should be accompanied by a warning to caution subsequent users about the high levels of error, associated with the estimates.
3) Unacceptable	Estimates have a sample size of less than 30 (if CV tables are used), or very high coefficients of variation in excess of 33.3%. Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter F (or some similar identifier) and the following warning should accompany the estimates: “Please be warned that these estimates [flagged with the letter F] do not meet Statistics Canada’s quality standards. Conclusions based on these data will be unreliable, and most likely invalid.”

9.6 Release Cut-off's for the Maternity Experiences Survey

The following table provides an indication of the precision of population estimates as it shows the release cut-offs associated with each of the three quality levels presented in the previous section. These cut-offs are derived from the coefficient of variation (CV) tables discussed in Chapter 10.0.

For example, the table shows that the quality of a weighted estimate of 50 people possessing a given characteristic in Newfoundland and Labrador is marginal.

Note that these cut-offs apply to estimates of population totals only. To estimate ratios, users should not use the numerator value (nor the denominator) in order to find the corresponding quality level. Rule 4 in Section 10.1 and Example 4 in Section 10.1.1 explain the correct procedure to be used for ratios.

Provinces and Territories	**Acceptable CV 0.0% to 16.5%**	**Marginal CV 16.6% to 33.3%**	**Unacceptable CV > 33.3%**
Newfoundland and Labrador	90 & over	30 to < 90	under 30
Prince Edward Island	30 & over		under 30
Nova Scotia	136 & over	35 to < 136	under 35
New Brunswick	136 & over	36 to < 136	under 36
Quebec	500 & over	125 to < 500	under 125
Ontario	556 & over	138 to < 556	under 138
Manitoba	231 & over	61 to < 231	under 61
Saskatchewan	211 & over	55 to < 211	under 55
Alberta	475 & over	121 to < 475	under 121
British Columbia	471 & over	120 to < 471	under 120
Yukon	30 & over		under 30
Northwest Territories	30 & over		under 30
Nunavut	30 & over		under 30
Canada	**478 & over**	**118 to < 478**	**under 118**

10.0 Approximate Sampling Variability Tables

In order to supply coefficients of variation (CV) which would be applicable to a wide variety of categorical estimates produced from this microdata file and which could be readily accessed by the user, a set of Approximate Sampling Variability Tables has been produced. These CV tables allow the user to obtain an approximate coefficient of variation based on the size of the estimate calculated from the survey data.

The coefficients of variation are derived using the variance formula for simple random sampling and incorporating a factor which reflects the nature of the sample design. This factor, known as the design effect, was determined by first calculating design effects for a wide range of characteristics and then choosing from among these a conservative value (usually the 75th percentile) to be used in the CV tables which would then apply to the entire set of characteristics.

The table below shows the conservative value of the design effects as well as sample sizes and population counts by province which were used to produce the Approximate Sampling Variability Tables for the Maternity Experiences Survey (MES).

Provinces and Territories	Design Effect	Sample Size	Population
Newfoundland and Labrador	1.01	279	1,020
Prince Edward Island	0.92	184	288
Nova Scotia	1.00	344	1,725
New Brunswick	1.04	303	1,488
Quebec	1.03	1,256	18,333
Ontario	1.03	1,858	29,688
Manitoba	1.00	341	2,692
Saskatchewan	1.02	341	2,443
Alberta	1.01	651	9,435
British Columbia	1.02	631	8,997
Yukon	0.88	61	82
Northwest Territories	0.78	89	146
Nunavut	0.92	83	170
Canada	**1.20**	**6,421**	**76,508**

All coefficients of variation in the Approximate Sampling Variability Tables are approximate and, therefore, unofficial. Estimates of actual variance for specific variables may be obtained from Statistics Canada on a cost-recovery basis or by using the bootstrap weights provided with the data. Since the approximate CV is conservative, the use of actual variance estimates may cause the estimate to be switched from one quality level to another. For instance a *marginal* estimate could become *acceptable* based on the exact CV calculation.

Remember: If the number of observations on which an estimate is based is less than 30, the weighted estimate is most likely unacceptable and Statistics Canada recommends not to release such an estimate, regardless of the value of the coefficient of variation.

10.1 How to Use the Coefficient of Variation Tables for Categorical Estimates

The following rules should enable the user to determine the approximate coefficients of variation from the Approximate Sampling Variability Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between such estimates.

Rule 1: Estimates of Numbers of Persons Possessing a Characteristic (Aggregates)

The coefficient of variation depends only on the size of the estimate itself. On the Approximate Sampling Variability Table for the appropriate geographic area, locate the estimated number in the left-most column of the table (headed "Numerator of Percentage") and follow the asterisks (if any) across to the first figure encountered. This figure is the approximate coefficient of variation.

Rule 2: Estimates of Proportions or Percentages of Persons Possessing a Characteristic

The coefficient of variation of an estimated proportion or percentage depends on both the size of the proportion or percentage and the size of the total upon which the proportion or percentage is based. Estimated proportions or percentages are relatively more reliable than the corresponding estimates of the numerator of the proportion or percentage, when the proportion or percentage is based upon a sub-group of the population. For example, the proportion of women whose primary caregiver at birth was a midwife and who rated their labour and birth experiences as "very positive", is more reliable than the estimated number of women whose primary caregiver at birth was a midwife and who rated their labour and birth experiences as "very positive". (Note that in the tables the coefficients of variation decline in value reading from left to right).

When the proportion or percentage is based upon the total population of the geographic area covered by the table, the CV of the proportion or percentage is the same as the CV of the numerator of the proportion or percentage. In this case, Rule 1 can be used.

When the proportion or percentage is based upon a subset of the total population (e.g. those in a particular age group), reference should be made to the proportion or percentage (across the top of the table) and to the numerator of the proportion or percentage (down the left side of the table). The intersection of the appropriate row and column gives the coefficient of variation.

Rule 3: Estimates of Differences Between Aggregates or Percentages

The standard error of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard error considered separately. That is, the standard error of a difference $\left(\hat{d} = \hat{X}_1 - \hat{X}_2\right)$ is:

$$\sigma_{\hat{d}} \sqrt{\left(\hat{X}_1 \alpha_1\right)^2 + \left(\hat{X}_2 \alpha_2\right)^2}$$

where $\hat{X}_1$ is estimate 1, $\hat{X}_2$ is estimate 2, and α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d}$. This formula is accurate for the difference between separate and uncorrelated characteristics, but is only approximate otherwise.

Rule 4: Estimates of Ratios

In the case where the numerator is a subset of the denominator, the ratio should be converted to a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is the number of women whose primary caregiver at birth was a midwife, and the numerator is the number of women whose primary caregiver at birth was a midwife and who rated their labour and birth experiences as "very positive".

In the case where the numerator is not a subset of the denominator, as for example, the ratio of the number of baby boys admitted to an intensive care or special care unit immediately after birth, as compared to the number of baby girls admitted to an intensive care or special care unit immediately after birth, the standard error of the ratio of the estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately multiplied by $\hat{R}$. That is, the standard error of a ratio $\left(\hat{R} = \hat{X}_1 / \hat{X}_2\right)$ is:

$$\sigma_{\hat{R}} = \hat{R}\sqrt{\alpha_1^{\ 2} + \alpha_2^{\ 2}}$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The coefficient of variation of $\hat{R}$ is given by $\sigma_{\hat{R}} / \hat{R}$. The formula will tend to overstate the error if $\hat{X}_1$ and $\hat{X}_2$ are positively correlated and understate the error if $\hat{X}_1$ and $\hat{X}_2$ are negatively correlated.

Rule 5: Estimates of Differences of Ratios

In this case, Rules 3 and 4 are combined. The CVs for the two ratios are first determined using Rule 4, and then the CV of their difference is found using Rule 3.

10.1.1 Examples of Using the Coefficient of Variation Tables for Categorical Estimates

The following examples based on the MES are included to assist users in applying the foregoing rules.

Example 1: Estimates of Numbers of Persons Possessing a Characteristic (Aggregates)

Suppose that a user estimates that midwives were the primary caregiver at births for 3,207 women. How does the user determine the coefficient of variation of this estimate?

1) Refer to the coefficient of variation table for CANADA.

2) The estimated aggregate 3,207 does not appear in the left-hand column (the "Numerator of Percentage" column), so it is necessary to use the figure closest to it, namely 3,000.

3) The coefficient of variation for an estimated aggregate is found by referring to the first non-asterisk entry on that row, namely, 6.4%.

4) So the approximate coefficient of variation of the estimate is 6.4%. The finding that midwives were the primary caregiver at births for 3,207 women (to be rounded according to the rounding guidelines in Section 9.1) in the reference period is publishable with no qualifications.

Example 2: Estimates of Proportions or Percentages of Persons Possessing a Characteristic

Suppose that the user estimates that 2,268 / 3,207 = 70.7% of women whose primary caregiver at birth was a midwife, rated their labour and birth experiences as "very positive". How does the user determine the coefficient of variation of this estimate?

1) Refer to the coefficient of variation table for CANADA.

2) Because the estimate is a percentage which is based on a subset of the total population (i.e., women whose primary caregiver at birth was a midwife), it is necessary to use both the percentage (70.7%) and the numerator portion of the percentage (2,268) in determining the coefficient of variation.

3) The numerator, 2,268, does not appear in the left-hand column (the "Numerator of Percentage" column) so it is necessary to use the figure closest to it, namely 2,500. Similarly, the percentage estimate does not appear as any of the column headings, so it is necessary to use the percentage closest to it, 70.0%.

4) The figure at the intersection of the row and column used, namely 4.0% is the coefficient of variation to be used.

5) So the approximate coefficient of variation of the estimate is 4.0%. The finding that 70.7% of women whose primary caregiver at birth was a midwife, rated their labour and birth experiences as "very positive" can be published with no qualifications.

Example 3: Estimates of Differences Between Aggregates or Percentages

Suppose that a user estimates that the primary caregiver at birth for 2,268 / 3,207 = 70.7% of women was a midwife and for 36,688 / 69,376 = 52.9% of women an obstetrician, gynaecologist, family doctor or nurse or nurse practitioner, and they rated their labour and birth experiences as "very positive". How does the user determine the coefficient of variation of the difference between these two estimates?

1) Using the CANADA coefficient of variation table in the same manner as described in Example 2 gives the CV of the estimate for women whose primary caregiver at birth was a midwife as 4.0%, and the CV of the estimate for women whose primary caregiver at birth was a obstetrician, gynaecologist, family doctor or nurse or nurse practitioner as 1.5%.

2) Using Rule 3, the standard error of a difference $\left(\hat{d} = \hat{X}_1 - \hat{X}_2\right)$ is:

$$\sigma_{\hat{d}} = \sqrt{\left(\hat{X}_1\alpha_1\right)^2 + \left(\hat{X}_2\alpha_2\right)^2}$$

where $\hat{X}_1$ is estimate 1 (women whose primary caregiver at birth was a midwife), $\hat{X}_2$ is estimate 2 (women whose primary caregiver at birth was an obstetrician, gynaecologist, family doctor or nurse or nurse practitioner), and α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively.

That is, the standard error of the difference $\hat{d} = 0.707 - 0.529 = 0.178$ is:

$$\sigma_{\hat{d}} = \sqrt{[(0.707)(0.040)]^2 + [(0.529)(0.015)]^2}$$
$$= \sqrt{(0.0007997) + (0.0000629)}$$
$$= 0.02937$$

3) The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d} = 0.02937 / 0.178 = 0.165$

4) So the approximate coefficient of variation of the difference between the estimates is 16.5%. The difference between the estimates is considered acceptable and this estimate can be released with no qualifications.

Example 4: Estimates of Ratios

Suppose that the user estimates that in British Columbia 646 baby boys and 295 baby girls were admitted to an intensive care or special care unit immediately after birth. The user is interested in comparing the estimate of baby boys versus that of baby girls in British Columbia in the form of a ratio. How does the user determine the coefficient of variation of this estimate?

1) First of all, this estimate is a ratio estimate, where the numerator of the estimate ($\hat{X}_1$) is the number of baby boys admitted to an intensive care or special care unit immediately after birth in British Columbia. The denominator of the estimate ($\hat{X}_2$) is the number of baby girls admitted to an intensive care or special care unit immediately after birth in British Columbia .

2) Refer to the coefficient of variation table for British Columbia.

3) The numerator of this ratio estimate is 646. The figure closest to it is 650. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 13.7%.

4) The denominator of this ratio estimate is 295. The figure closest to it is 300. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 20.7%.

5) So the approximate coefficient of variation of the ratio estimate is given by Rule 4, which is:

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^{\,2} + \alpha_2^{\,2}}$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. That is:

$$\alpha_{\hat{R}} = \sqrt{(0.137)^2 + (0.207)^2}$$
$$= \sqrt{0.0188 + 0.0428}$$
$$= 0.248$$

6) The obtained ratio of baby boys versus baby girls admitted to an intensive care or special care unit immediately after birth in British Columbia is 646 / 295 which is 2.19

(to be rounded according to the rounding guidelines in Section 9.1). The coefficient of variation of this estimate is 24.8%. The quality of the estimate is considered marginal. The estimate should be flagged with the letter E (or some similar identifier), and accompanied by a warning to caution subsequent users about the high level of error associated with the estimate.

Example 5: Estimates of Differences of Ratios

Suppose that the user estimates that the ratio of baby boys to baby girls admitted to an intensive care or special care unit immediately after birth is 1.30 for Ontario and 2.19 for British Columbia. The user is interested in comparing the two ratios to see if there is a statistical difference between them. How does the user determine the coefficient of variation of the difference?

1) First calculate the approximate coefficient of variation for the Ontario ratio ($\hat{R}_1$) and the British Columbia ratio ($\hat{R}_2$) as in Example 4. The approximate CV for the Ontario ratio is 12.7% and 24.8% for British Columbia.

2) Using Rule 3, the standard error of a difference ($\hat{d} = \hat{R}_1 - \hat{R}_2$) is:

$$\sigma_{\hat{d}} = \sqrt{\left(\hat{R}_1\alpha_1\right)^2 + \left(\hat{R}_2\alpha_2\right)^2}$$

where α_1 and α_2 are the coefficients of variation of $\hat{R}_1$ and $\hat{R}_2$ respectively. That is, the standard error of the difference $\hat{d}$ = 1.30 – 2.19 = -0.89 is:

$$\begin{aligned}\sigma_{\hat{d}} &= \sqrt{\left[(1.30)(0.127)\right]^2 + \left[(2.19)(0.248)\right]^2} \\ &= \sqrt{(0.0273) + (0.2950)} \\ &= 0.568\end{aligned}$$

3) The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d}$ = 0.568 / (-0.89) = -0.638

4) So the approximate coefficient of variation of the difference between the estimates is 63.8%. The difference between the estimates is considered unacceptable and Statistics Canada recommends this estimate not be released. However, should the user choose to do so, the estimate should be flagged with the letter F (or some similar identifier) and be accompanied by a warning to caution subsequent users about the high levels of error, associated with the estimate.

10.2 How to Use the Coefficient of Variation Tables to Obtain Confidence Limits

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example a 95% confidence interval can be described as follows:

If sampling of the population is repeated indefinitely, each sample leading to a new confidence interval for an estimate, then in 95% of the samples the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about 99 out of 100 that the difference would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate, $\hat{X}$, are generally expressed as two numbers, one below the estimate and one above the estimate, as $\left(\hat{X} - k, \hat{X} + k\right)$ where k is determined depending upon the level of confidence desired and the sampling error of the estimate.

Confidence intervals for an estimate can be calculated directly from the Approximate Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate $\hat{X}$, and then using the following formula to convert to a confidence interval ($CI_{\hat{x}}$):

$$CI_{\hat{x}} = \left(\hat{X} - t\hat{X}\alpha_{\hat{x}}, \hat{X} + t\hat{X}\alpha_{\hat{x}}\right)$$

where $\alpha_{\hat{x}}$ is the determined coefficient of variation of $\hat{X}$, and

$t = 1$ if a 68% confidence interval is desired;
$t = 1.6$ if a 90% confidence interval is desired;
$t = 2$ if a 95% confidence interval is desired;
$t = 2.6$ if a 99% confidence interval is desired.

Note: Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is not releasable, then the confidence interval is not releasable either.

10.2.1 Example of Using the Coefficient of Variation Tables to Obtain Confidence Limits

A 95% confidence interval for the estimated proportion of women whose primary caregiver at birth was a midwife and who rated their labour and birth experiences as “very positive” (from Example 2, Section 10.1.1) would be calculated as follows:

$\hat{X}$ = 70.7% (or expressed as a proportion 0.707)

t = 2

$\alpha_{\hat{x}}$ = 4.0% (0.040 expressed as a proportion) is the coefficient of variation of this estimate as determined from the tables.

$CI_{\hat{x}}$ = {0.707 – (2) (0.707) (0.040), 0.707 + (2) (0.707) (0.040)}

$CI_{\hat{x}}$ = {0.707 – 0.057, 0.707 + 0.057}

$CI_{\hat{x}}$ = {0.650, 0.764}

With 95% confidence it can be said that between 65.0% and 76.4% of women whose primary caregiver at birth was a midwife rated their labour and birth experiences as "very positive".

10.3 How to Use the Coefficient of Variation Tables to Do a T-test

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let $\hat{X}_1$ and $\hat{X}_2$ be sample estimates for two characteristics of interest. Let the standard error on the difference $\hat{X}_1 - \hat{X}_2$ be $\sigma_{\hat{d}}$.

If $t = \dfrac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}}$ is between -2 and 2, then no conclusion about the difference between the characteristics is justified at the 5% level of significance. If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 0.05 level. That is to say that the difference between the estimates is significant.

10.3.1 Example of Using the Coefficient of Variation Tables to Do a T-test.

Let us suppose that the user wishes to test, at 5% level of significance, the hypothesis that there is no difference between the proportion of women whose primary caregiver at birth was a midwife and the proportion of women whose primary caregiver at birth was an obstetrician, gynaecologist, family doctor or nurse or nurse practitioner, and who rated their labour and birth experiences as "very positive". From Example 3, Section 10.1.1, the standard error of the difference between these two estimates was found to be 0.02937. Hence,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}} = \frac{0.707 - 0.529}{0.02937} = \frac{0.178}{0.02937} = 6.06$$

Since t = 6.06 is greater than 2, it must be concluded that there is a significant difference between the two estimates at the 0.05 level of significance.

10.4 Coefficients of Variation for Quantitative Estimates

For quantitative estimates, special tables would have to be produced to determine their sampling error. Since most of the variables for the MES are primarily categorical in nature, this has not been done.

As a general rule, however, the coefficient of variation of a quantitative total will be larger than the coefficient of variation of the corresponding category estimate (i.e., the estimate of the number of persons contributing to the quantitative estimate). If the corresponding category estimate is not releasable, the quantitative estimate will not be either. For example, the coefficient of variation of the total number of prenatal care visits would be greater than the coefficient of variation of the corresponding proportion of mothers who received prenatal care. Hence, if the coefficient of variation of the proportion is unacceptable (making the proportion not releasable), then the coefficient of variation of the corresponding quantitative estimate will also be unacceptable (making the quantitative estimate not releasable).

Coefficients of variation of such estimates can be derived using the bootstrap weights provided with the data (see Section 9.4).

10.5 Coefficient of Variation Tables

Refer to MES2006_CVTabsE.pdf for the coefficient of variation tables.

11.0 Weighting

This chapter outlines the weighting steps that were performed to derive the final weights for the Maternity Experiences Survey (MES). The records were first categorized as follows:

Category		Description	Count
Unresolved (UNR)	UNR1	Not sent to the field for data collection (e.g., due to lack of contact information)	70
	UNR2	Sent to the field, but no data received to determine whether in-scope for the survey.	1,285
	UNR3	Partial data received, but not sufficient to determine whether in-scope for the survey (in-scope according to SRC module, VSB module incomplete).	250
Out-of-scope (OOS)	OOS1	Out-of-scope Type 1 (OOS according to SRC module).	99
	OOS2	Out-of-scope Type 2 (OOS according to VSB module).	135
In-scope	RESP	Respondents (RESP): Sufficient data and permission to share was given.	6,421
	NR	Non-respondents (NR): Insufficient data or no permission to share data was given.	282
		Total initial sample size	**8,542**

Note: The SRC module had questions regarding the presence of the baby in the household (at least one night per month), the name of the mother and her presence in the household. If the wrong person was identified as the mother on the sample file, the interviewer asked to speak to the mother.

The VSB module asked questions about the mother and baby's date of birth, verified that the mother was the birth mother, confirmed that the baby was born in Canada, and confirmed that the baby was a single birth.

Whenever there were sufficient numbers of units, the weighting adjustments were calculated within weighting classes which, for the most part, corresponded to our original strata. Additional weighting classes were defined in New Brunswick, Quebec and British Columbia for older mothers based on first language and for older mothers in the territories based on Aboriginal status. The following steps were performed:

1. Calculate design weights

Each of the 8,542 records in the initial sample was assigned a design weight, W_1, equal to the census weight multiplied by the inverse of the probability of selection within each stratum, as follows:

$$W_1 = \text{Census weight} \times \left(\frac{\text{Number of units on the census frame}}{\text{Number of units in MES sample}} \right)$$

The census weight for records in the provinces was the number of Form 2A, Form 2B and Form 2D questionnaires divided by the number of Form 2A questionnaires (approximately 5/4) within the collection unit (CU); for records in the North, the census weights were one.

2. Adjust for units not sent to the field (UNR1)

Weighting adjustments were performed to take into account 70 records that were not sent to the field for data collection. The records were grouped into weighting classes, and the following adjustments were calculated within each weighting class:

$$W_2 = \left(\frac{\sum W_1 \text{ for records sent to the field} + \sum W_1 \text{ for records not sent to the field}}{\sum W_1 \text{ for records sent to the field}} \right) \times W_1 .$$

3. Adjust for unresolved units with no data (UNR2)

Weighting adjustments were performed to take into account 1,285 records that were sent to the field for data collection but for which no data were received, to determine whether they were in-scope for the survey. The records were grouped into weighting classes, and the following adjustments were calculated within each weighting class:

$$W_3 = \left(\frac{\sum W_2 \text{ for records classified as RESP, NR, OOS1, OOS2, UNR3} + \sum W_2 \text{ for records classified as UNR2}}{\sum W_2 \text{ for records classified as RESP, NR, OOS1, OOS2, UNR3}} \right) \times W_2$$

4. Adjust for unresolved units with partial data (UNR3)

Weighting adjustments were performed to take into account 250 records for which partial data, but not sufficient data were collected to determine whether they are in-scope for the survey. The records were grouped into weighting classes, and the following adjustments were calculated within each weighting class:

$$W_4 = \left(\frac{\sum W_3 \text{ for records classified as RESP, NR, OOS2} + \sum W_3 \text{ for records classified as UNR3}}{\sum W_3 \text{ for records classified as RESP, NR, OOS2}} \right) \times W_3$$

5. Adjust for non-respondents (NR)

Weighting adjustments were performed to take into account 282 in-scope non-respondents. These are records with insufficient data or the respondent did not agree to share their data with the Public Health Agency of Canada. The records were grouped into weighting classes, and the following adjustments were calculated within each weighting class:

$$W_5 = \left(\frac{\sum W_4 \text{ for records classified as RESP} + \sum W_4 \text{ for records classified as NR}}{\sum W_4 \text{ for records classified as RESP}} \right) \times W_4$$

6. Calibration

The original MES frame used unedited Census data and did not fully cover the target population. The MES weights were calibrated to provincial / territorial estimates of the survey population based on the final Census data as follows:

$$W_6 = \left(\frac{\text{Final Census Counts}}{\sum W_5 \text{ for records classified as RESP} + \sum W_5 \text{ for some OOS records}} \right) \times W_5$$

The out-of-scope records included in the denominator are mothers who gave birth outside Canada, non-birth mothers, deceased or institutionalized mothers or babies, and mothers who no longer live with their babies. This takes into consideration the expected overcoverage of the survey population by the final Census counts (which is not known) in accordance with the weighted representation of these units in our sample (which is known).

12.0 Questionnaire

Refer to the file MES2006_QuestE.pdf for the English questionnaire used to collect the data.

13.0 Record Layout with Univariate Frequencies

See MES2006_Share_CdBk.pdf for the record layout with univariate counts.