

**Canadian Community Health Survey  
Cycle 2.1, 2003  
Dummy Files for Remote Access to the Master Files**

June 2005

*Notice*

*The dummy files should not be used for purposes other than to develop and test the computer programs that are to be submitted by remote access. The dummy files contain modified data, and must never be used for analytical purposes.*

**1. Background and Overview**

The Canadian Community Health Survey (CCHS) is a cross-sectional survey that collects information related to health status, health care utilization and health determinants for the Canadian population. The CCHS operates on a two-year collection cycle. The first year of the survey cycle “.1” is a large sample, general population health survey, designed to provide reliable estimates at the health region level. The second year of the survey cycle “.2” is a smaller survey designed to provide provincial level results on specific focused health topics.

In order to provide extensive access to the data, a public use microdata file (PUMF) was produced for Cycle 2.1. The creation of a PUMF involves the application of rigorous procedures to ensure data confidentiality. Consequently, survey variables may have to be grouped, capped, or simply suppressed.

As an alternative, in cases where the PUMF does not meet their needs, researchers can perform their analyses on the CCHS master file(s). In order to provide researchers with a means to access the master file(s), a remote access facility has been implemented. Remote access provides researchers with the possibility to submit computer programs via e-mail to a dedicated address ([cchs-esc@statcan.ca](mailto:cchs-esc@statcan.ca)), and to receive the results by return e-mail. To obtain remote access privileges, it is necessary that researchers obtain advance approval from the Health Statistics Division. Requests must be submitted to the aforementioned e-mail address and must provide the following, clearly itemized information:

- the researcher’s affiliation,
- the name of all researchers involved in the project,
- the title of the research project,
- an abstract of the project,
- the goals of the research,
- the data to which access is required (survey, cycle),
- why the project requires access to the master data rather than the PUMF,
- why Remote Access service is chosen rather the on-site access in a Research Data Centre (RDC),
- the expected results, and
- the project’s expected completion date.

Further information is available by contacting the CCHS team at the above e-mail address or by phone at (613) 951-1653. Once the request for remote access has been approved, the researcher can submit his/her computer programs to the CCHS team for processing on the master file(s). The computer output is reviewed by the team for confidentiality concerns and returned to the researcher. However, the correctness and accuracy of each program submission remains, at all times, the sole responsibility of the researcher.

With the dummy files supplied on this CD-ROM, the researcher can develop and test his/her computer programs before submitting them to the CCHS team. For CCHS 2.1, the five following dummy data files were created:

- **Full sample file:** File with all respondents, and all variables from common and optional modules; 48,078 records.
- **Subsample 1 file:** File with only respondents who were administered subsample 1 modules. Variables included are the ones specifically collected for subsample 1, as well as the ones from the common modules; 14,477 records.
- **Subsample 2 file:** Similar to subsample 1 file but specific to subsample 2; 29,474 records.
- **Subsample 3 file:** Similar to subsample 1 file but specific to subsample 3; 14,241 records.
- **Group file:** File with all respondents, but only grouped variables (from the full sample file) are included.

While certain administrative variables (which are of no analytical interest) have been recoded to “blank” or “9”, the dummy files simulate the master files perfectly. They contain the same variables and have the same record layout. The data values, however, have been modified in order to protect the confidentiality of respondents. However, note that the number of records has been reduced by more than 60% compared to what is available in master versions of these files.

Corresponding dummy bootstrap weights files were also created for each of these files (except for the Group file which consists of the same sample as the full sample file, therefore, bootstrap weights for that later file must be used when dealing with the Group file). The CD-ROM includes SAS and SPSS macro programs (Bootvar) for calculating the variance of estimates using these bootstrap weights.

The CD-ROM also contains peripheral SAS and SPSS layout statement files, as well as the technical and methodological documentation usually accompanying the master files. Please refer to Section 3 of the present document for the complete list of files on this CD-ROM.

The following sections of this document describe in more detail the steps leading to the creation of the dummy files.

## **2. Creation of the dummy files**

### **2.1 Classes of Variables**

All records are first classified by age, sex. The purpose for this is to create classes of records having the same questionnaire flow so that, after random permutation of the data within the classes, each dummy record contains data respecting a certain internal consistency.

### **2.2 Sampling and Restrictions of Variable Values**

As an additional measure of preserving data confidentiality in the master files, the dummy files are created by sub-sampling the master files records. Also, quantitative variables are top-coded and bottom-coded. Variables considered as sensitive or without analytical value are recoded to “blank” or to “9” (see section 2.4). All of these modifications mean that we do not necessarily find, on the dummy files, all the possible categories for each variable. To obtain complete information on a specific variable, users are asked to consult the master files documentation. Finally, all the weight variables have been repoststratified according to the same methods used to poststratify the master files weights. However, it should be noted that the weights remain “dummy” weights. Other details related to the creation of this file are not available because they would disclose too much information on the method used for creating the dummy data, hence putting confidentiality at risk.

### **2.3 Block of variables**

The third step consists of creating blocks of variables. Dummy records are produced by randomly swapping blocks of variables for the records in the same class. At first, elementary blocks of variables are grouped following the various modules of the questionnaire. Blocks do not always correspond to modules of the questionnaire, however, as some modules are split differently. If the content of the elementary block depends on the answers obtained in a previous block, the respective elementary block is amalgamated with the other one to preserve the internal consistency of the records.

## 2.4 Recode Variables to “blank” and to “9”

The following variables are recoded to “blank” or to “9” during the creation of the dummy files:

PERSONID SAMC\_TYP SAMC\_CP SAMCDCP2 SAMCDSHR SAMCDSJB SAMCDLNK ADMC\_STA ADMC\_PRX  
ADMC\_YOI AMDC\_MOI ADMCDOI ADMCFQCP ADMC\_YQI ADMC\_MQI ADMC\_DQI ADMC\_N09 ADMC\_N10  
ADMC\_N11 ADMC\_LHH ADMC\_N12 DHHC\_YOM DHHC\_MOB DHHC\_DOB

## 3. Content of CD-ROM

LISEZ_MOI.PDF	French documentation about remote-access, creation of the Dummy File and contents of the CD-ROM
READ_ME.PDF	English documentation about remote-access, creation of the Dummy File and contents of the CD-ROM

*Dummy File :*

DATA\ HS.TXT	CCHS 2.1 dummy file for the full sample (48,078 records)
HS_S1.TXT	CCHS 2.1 dummy file for subsample 1 (14,477 records)
HS_S2.TXT	CCHS 2.1 dummy file for subsample 2 (29,474 records)
HS_S3.TXT	CCHS 2.1 dummy file for subsample 3 (14,241 records)
HSG.TXT	CCHS 2.1 dummy file for the grouped variables file (48,078 records)

*Record Layouts, Statements :*

LAYOUT\ HS_fmt.sas	SAS Format Statement.
HS_i.sas	SAS Infile and Input Statements.
HS_lbe.sas	SAS Label Statement in English.
HS_lbf.sas	SAS Label Statement in French.
HS_o.sas	SAS Output Statement.
HS_pfe.sas	SAS Proc Format Statement in English.
HS_pff.sas	SAS Proc Format Statement in French.
HSG_fmt.sas	Grouped SAS Format Statement.
HSG_i.sas	Grouped SAS Infile and Input Statements.
HSG_lbe.sas	Grouped SAS Label Statement in English.
HSG_lbf.sas	Grouped SAS Label Statement in French.
HSG_o.sas	Grouped SAS Output Statement.
HSG_pfe.sas	Grouped SAS Proc Format Statement in English.
HSG_pff.sas	Grouped SAS Proc Format Statement in French.
HSS1_fmt.sas	Sub-sample 1 SAS Format Statement.
HSS1_i.sas	Sub-sample 1 SAS Infile and Input Statements.
HSS1_lbe.sas	Sub-sample 1 SAS Label Statement in English.
HSS1_lbf.sas	Sub-sample 1 SAS Label Statement in French.
HSS1_o.sas	Sub-sample 1 SAS Output Statement.
HSS1_pfe.sas	Sub-sample 1 SAS Proc Format Statement in English.
HSS1_pff.sas	Sub-sample 1 SAS Proc Format Statement in French.
HSS2_fmt.sas	Sub-sample 2 SAS Format Statement.
HSS2_i.sas	Sub-sample 2 SAS Infile and Input Statements.
HSS2_lbe.sas	Sub-sample 2 SAS Label Statement in English.

HSS2_lbf.sas	Sub-sample 2 SAS Label Statement in French.
HSS2_o.sas	Sub-sample 2 SAS Output Statement.
HSS2_pfe.sas	Sub-sample 2 SAS Proc Format Statement in English.
HSS2_pff.sas	Sub-sample 2 SAS Proc Format Statement in French.
HSS3_fmt.sas	Sub-sample 3 SAS Format Statement.
HSS3_i.sas	Sub-sample 3 SAS Infile and Input Statements.
HSS3_lbe.sas	Sub-sample 3 SAS Label Statement in English.
HSS3_lbf.sas	Sub-sample 3 SAS Label Statement in French.
HSS3_o.sas	Sub-sample 3 SAS Output Statement.
HSS3_pfe.sas	Sub-sample 3 SAS Proc Format Statement in English.
HSS3_pff.sas	Sub-sample 3 SAS Proc Format Statement in French.
HS_i.sps	SPSS Infile Statement.
HSmiss.sps	SPSS Missing Values Statement.
HSvale.sps	SPSS Value Labels Statement in English.
HSvalf.sps	SPSS Value Labels Statement in French.
HSvare.sps	SPSS Variable Labels Statement in English.
HSvarf.sps	SPSS Variable Labels Statement in French.
HSG_i.sps	Grouped SPSS Infile Statement.
HSGmiss.sps	Grouped SPSS Missing Values Statement.
HSGvale.sps	Grouped SPSS Value Labels Statement in English.
HSGvalf.sps	Grouped SPSS Value Labels Statement in French.
HSGvare.sps	Grouped SPSS Variable Labels Statement in English.
HSGvarf.sps	Grouped SPSS Variable Labels Statement in French.
HSS1_i.sps	Sub-sample 1 SPSS Infile Statement.
HSS1miss.sps	Sub-sample 1 SPSS Missing Values Statement.
HSS1vale.sps	Sub-sample 1 SPSS Value Labels Statement in English.
HSS1valf.sps	Sub-sample 1 SPSS Value Labels Statement in French.
HSS1vare.sps	Sub-sample 1 SPSS Variable Labels Statement in English.
HSS1varf.sps	Sub-sample 1 SPSS Variable Labels Statement in French.
HSS2_i.sps	Sub-sample 2 SPSS Infile Statement.
HSS2miss.sps	Sub-sample 2 SPSS Missing Values Statement.
HSS2vale.sps	Sub-sample 2 SPSS Value Labels Statement in English.
HSS2valf.sps	Sub-sample 2 SPSS Value Labels Statement in French.
HSS2vare.sps	Sub-sample 2 SPSS Variable Labels Statement in English.
HSS2varf.sps	Sub-sample 2 SPSS Variable Labels Statement in French.
HSS3_i.sps	Sub-sample 3 SPSS Infile Statement.
HSS3miss.sps	Sub-sample 3 SPSS Missing Values Statement.
HSS3vale.sps	Sub-sample 3 SPSS Value Labels Statement in English.
HSS3valf.sps	Sub-sample 3 SPSS Value Labels Statement in French.
HSS3vare.sps	Sub-sample 3 SPSS Variable Labels Statement in English.
HSS3varf.sps	Sub-sample 3 SPSS Variable Labels Statement in French.
readfile.sps	Program to read, label and format the ASCII format data file into SPSS format.
readfile.sas	Program to read, label and format the ASCII format data file into SAS format.

*Documentation:*

DOC\

English DvDoc.pdf	Derived Variable Documentation in English.
English Questionnaire.pdf	Questionnaire in English.
English Household Questionnaire.pdf	Household Questionnaire in English.
English Data Dictionary (Freqs).pdf	Data Dictionary in English.
English DD Alpha Index.pdf	Alphabetical Index in English.
English DD Topical Index.pdf	Topical Index in English.

English Record Layout.pdf	Record Layout in English.
French DvDoc.pdf	Derived Variable Documentation in French.
French Questionnaire.pdf	Questionnaire in French.
French Household Questionnaire.pdf	Household Questionnaire in French.
French Data Dictionary (Freqs).pdf	Data Dictionary in French.
French DD Alpha Index.pdf	Alphabetical Index in French.
French DD Topical Index.pdf	Topical Index in French.
French Record Layout.pdf	Record Layout in French.
Modestudy.pdf	Documentation on the comparison of CAPI and CATI in English.
Usage des sous-échantillons - fichier partagé.pdf	Documentation on the sub-samples in French.
Use of sub-samples - share.pdf	Documentation on the sub-samples in English.

#### Group\

English Data Dictionary (Freqs).pdf	Grouped Data Dictionary in English.
English DD Alpha Index.pdf	Grouped Alphabetical Index in English.
English DD Topical Index.pdf	Grouped Topical Index in English.
English Record Layout.pdf	Grouped Record Layout in English.
French Data Dictionary (Freqs).pdf	Grouped Data Dictionary in French.
French DD Alpha Index.pdf	Grouped Alphabetical Index in French.
French DD Topical Index.pdf	Grouped Topical Index in French.
French Record Layout.pdf	Grouped Record Layout in French.

#### Sub-sample 1\

English Data Dictionary (Freqs).pdf	Sub-sample 1 Data Dictionary in English.
English DD Alpha Index.pdf	Sub-sample 1 Alphabetical Index in English.
English DD Topical Index.pdf	Sub-sample 1 Topical Index in English.
English Record Layout.pdf	Sub-sample 1 Record Layout in English.
French Data Dictionary (Freqs).pdf	Sub-sample 1 Data Dictionary in French.
French DD Alpha Index.pdf	Sub-sample 1 Alphabetical Index in French.
French DD Topical Index.pdf	Sub-sample 1 Topical Index in French.
French Record Layout.pdf	Sub-sample 1 Record Layout in French.

#### Sub-sample 2\

English Data Dictionary (Freqs).pdf	Sub-sample 2 Data Dictionary in English.
English DD Alpha Index.pdf	Sub-sample 2 Alphabetical Index in English.
English DD Topical Index.pdf	Sub-sample 2 Topical Index in English.
English Record Layout.pdf	Sub-sample 2 Record Layout in English.
French Data Dictionary (Freqs).pdf	Sub-sample 2 Data Dictionary in French.
French DD Alpha Index.pdf	Sub-sample 2 Alphabetical Index in French.
French DD Topical Index.pdf	Sub-sample 2 Topical Index in French.
French Record Layout.pdf	Sub-sample 2 Record Layout in French.

#### Sub-sample 3\

English Data Dictionary (Freqs).pdf	Sub-sample 3 Data Dictionary in English.
English DD Alpha Index.pdf	Sub-sample 3 Alphabetical Index in English.
English DD Topical Index.pdf	Sub-sample 3 Topical Index in English.
English Record Layout.pdf	Sub-sample 3 Record Layout in English.
French Data Dictionary (Freqs).pdf	Sub-sample 3 Data Dictionary in French.
French DD Alpha Index.pdf	Sub-sample 3 Alphabetical Index in French.
French DD Topical Index.pdf	Sub-sample 3 Topical Index in French.
French Record Layout.pdf	Sub-sample 3 Record Layout in French.

*Files related to bootstrap weights for variance calculation:*

#### BOOTSTRP\

##### DATA\

B5.TXT	Bootstrap weights for the full sample in ASCII format.
B5_S1.TXT	Bootstrap weights for the Subsample 1 file in ASCII format.
B5_S2.TXT	Bootstrap weights for the Subsample 2 file in ASCII format.
B5_S3.TXT	Bootstrap weights for the Subsample 3 file in ASCII format.

## SAS\

### DOC\

SASBootdoc_eng.pdf	English documentation of the BOOTVARE_V30.SAS program.
SASBootdoc_fr.pdf	French documentation of the BOOTVARF_V30.SAS program.
AppendixC_Health.pdf	Variable and file names to use with the BOOTVAR program in English.
AnnexeC_Santé.pdf	Variable and file names to use with the BOOTVAR program in French.

### PGM\

Bootvare_v30.sas	SAS bootstrap variance program with English comments.
Bootvarf_v30.sas	SAS bootstrap variance program with French comments.
Macro_e_v30.sas	Macro SAS program with English comments.
Macrof_v30.sas	Macro SAS program with French comments.
Step1.sas	Example, with English comments, of the SAS program creating the data analysis file.
Etape1.sas	Example, with French comments, of the SAS program creating the data analysis file.

## SPSS\

### DOC\

SPSSBootdoc_eng.pdf	English documentation of the BOOTVARE_V30.SPS program.
SPSSBootdoc_fr.pdf	French documentation of the BOOTVARF_V30.SPS program.
AppendixC_Health.pdf	Variable and file names to use with the BOOTVAR program in English.
AnnexeC_Santé.pdf	Variable and file names to use with the BOOTVAR program in French.

### PGM\

Bootvare_v30.sps	SPSS bootstrap variance program with English comments.
Bootvarf_v30.sps	SPSS bootstrap variance program with French comments.
Macro_e_v30.sps	Macro SPSS program with English comments.
Macrof_v30.sps	Macro SPSS program with French comments.
Step1.sps	Example, with English comments, of the SPSS program creating the data analysis file.
Etape1.sps	Example, with French comments, of the SPSS program creating the data analysis file.

## LAYOUT\

B5_i.sas	SAS Record Layout for the bootstrap weights file B5.TXT
B5_i.sps	SPSS Record Layout for the bootstrap weights file B5.TXT
B5_S1_i.sas	SAS Record Layout for the bootstrap weights file B5_S1.TXT for sub-sample 1.
B5_S1_i.sps	SPSS Record Layout for the bootstrap weights file B5_S1.TXT for sub-sample 1.
B5_S2_i.sas	SAS Record Layout for the bootstrap weights file B5_S2.TXT for sub-sample 2.
B5_S2_i.sps	SPSS Record Layout for the bootstrap weights file B5_S2.TXT for sub-sample 2.
B5_S3_i.sas	SAS Record Layout for the bootstrap weights file B5_S3.TXT for sub-sample 3.
B5_S3_i.sps	SPSS Record Layout for the bootstrap weights file B5_S3.TXT for sub-sample 3.