Cree Health Survey 2003
Canadian Community Health Survey Cycle 2.1
Iiyiyiu Aschii

Survey methods

June 2008

Conseil Cri de la santé et des services sociaux de la Baie James
ᓂᑐᑯᔮᐤ ᑳᔮ ᐄᒥᐲ ᐋᐄᓐ ᐊᓇᓇᑲᒋᑕᑲᓂᐅᒡ
Cree Board of Health and Social Services of James Bay

Institut national de santé publique Québec

Canadian Community Health Survey, Cycle 2.1
Iiyiyiu Aschii, 2003

Survey methods

AUTHOR

Denis Hamel
Unité Connaissance-surveillance, direction Recherche, formation et développement
Institut national de santé publique du Québec

EDITING AND COORDINATION

Gilles Légaré, Marie-Anne Kimpton and Mélanie Anctil
Unité Connaissance-surveillance, direction Recherche, formation et développement
Institut national de santé publique du Québec

STATISTICAL ANALYSES

Denis Hamel
Unité Connaissance-surveillance, direction Recherche, formation et développement
Institut national de santé publique du Québec

WITH THE COLLABORATION OF

Jill Torrie, Yv Bonnier-Viger, Marcellin Gangbè, Elena Kuzmina et Pierre Lejeune
Public Health Department
Cree Board of Health and Social Services of James Bay

REVIEWER

Louis Rochette
Unité Connaissance-surveillance, direction Recherche, formation et développement
Institut national de santé publique du Québec

TRANSLATION

Axel Harvey

LAYOUT

Line Mailloux
Unité Connaissance-surveillance, direction Recherche, formation et développement
Institut national de santé publique du Québec

PUBLICATION

Cree Board of Health and Social Services of James Bay &
Institut national de santé publique du Québec

The complete text of this document is available on the CBHSSJB's Web site at www.creepublichealth.org and on the INSPQ's Web site at www.inspq.qc.ca

Photo: Jimmy Sam
Design: Katya Petrov

Document deposited on Santécom (http://www.santecom.qc.ca)
Legal deposit – 2nd trimester 2008
Bibliothèque et Archives nationales du Québec
Library and Archives Canada
ISBN: 978-2-550-52896-8 (PDF)

TABLE OF CONTENTS

FOREWORD

This publication presents the findings of a health survey carried out in 2003 among households of Iiyiyiu Aschii[1]. A similar survey had been undertaken in the region by Santé Québec in 1991 (Santé Québec, 1994). Ten years later, the Public Health Department of the Cree Board of Health and Social Services of James Bay (CBHSSJB) urgently required a new picture of its population's state of health. The purpose of the 2003 survey was to gather up-to-date information on the region's main health problems and related factors in order to improve the planning, administration, and evaluation of various social and health programs.

According to the 2001 Public Health Act (*Loi sur la santé publique*), Quebec's public health departments must periodically assess the health of their respective populations. Since 2000-2001, the province's socio-sanitary regions – with the exception of Iiyiyiu Aschii and Nunavik – have participated in the Canadian Community Health Survey (CCHS) conducted by Statistics Canada.

In 2003 the Public Health Department of Iiyiyiu Aschii decided to take part in this vast project, which was already under way across Canada, and initiated a CCHS-type survey on its own territory (Statistics Canada, 2003). Because the CBHSSJB Public Health Department is connected to the network of Quebec's Department of Health and Social Services (*Ministère de la santé et des services sociaux*, MSSS), it was able to enlist the expert assistance of the *Institut national de santé publique du Québec* (INSPQ) in coordinating the analysis of the results. Professionals drawn from Quebec's health care community and the Public Health Department of Iiyiyiu Aschii, as well as academic experts in the field, were given the task of drafting the publications. The analyses include results on various aspects of health affecting residents of Iiyiyiu Aschii and they also provide comparisons with 1991 data from the region and 2003 data from the rest of Quebec (Santé Québec, 1994; Statistics Canada, 2003). These analyses are relevant for everyone concerned with the health of Iiyiyiu Aschii residents (professionals, administrators, planners, and researchers).

Ten publications were produced as part of this survey:

- *Demographic and social characteristics of the population living in Iiyiyiu Aschii*
- *Food habits, physical activity and body weight*
- *Cigarette consumption*
- *Lifestyles related to alcohol consumption, drugs and gambling*
- *Preventive practices and changes for improving health*
- *Health status, life expectancy and limitation of activities*
- *Injuries and transportation safety*
- *Mental health*
- *Use and perceptions of health services*
- *Survey methods*

A final publication, *Survey highlights*, offers a rapid overall view of the health study's results.

Many people contributed to this study at every stage in its progress. Particularly deserving of mention are the roles played by Jill Elaine Torrie, Director of Specialized Services, and Yv Bonnier-Viger, Director of Public Health of the Cree Board, throughout the planning phase and during operations on the field. Above all, we wish to thank the Cree population for its remarkable level of collaboration.

INTRODUCTION

This publication summarizes the methodological basis of the Canadian Community Health Survey (CCHS) carried out in Cree territory. An account of the relevant methods is necessary for a proper understanding of the results presented in the various topical publications and how they were arrived at. We shall first explain how the context in which the Cree survey was undertaken led to the choice of the Canadian survey and to other decisions. The sample design and all other methodological aspects of this survey, such as data collection, weighting, data quality, and dissemination of results, will be explained subsequently.

1. CONTEXT OF THE SURVEY

The last major study on the health of the James Bay Cree was undertaken by Santé Québec in 1991 with a sample of 354 households (Santé Québec, 1994). This was patterned after a survey carried out in Quebec in 1987, and included such additional topics as physical measurements and a 24-hour nutrition recall. Other factors specifically related to Cree communities were added; however, not all the data acquired in this survey were subjected to specific analysis.

[1] Please note that the socio-sanitary region for the James Bay Cree Territory is referred to by its Cree name, Iiyiyiu Aschii, throughout this text.

More than ten years later, it became absolutely necessary for the Cree Board of Health and Social Services of James Bay (CBHSSJB) to obtain a new and reliable picture of its client population's state of health. This second large-scale survey, by providing information on the population's main health-related problems and their contributing factors, would lead to better planning, administration, and evaluation of the region's various health programmes.

In the spring of 2003 two main scenarios were considered for the survey:

- Simply to repeat the survey carried out by Santé Québec in 1991.
- To make it part of the Canadian Community Health Survey (CCHS), which was already under way throughout Canada in 2003.

The second alternative was chosen for reasons which are easy to understand: (1) data collection could begin without delay because CCHS Cycle 2.1 was already proceeding; (2) it would be possible to make comparisons with the data already gathered in the rest of Quebec; (3) costs would be relatively low thanks to resources made available by Statistics Canada (a validated questionnaire, teams of interviewers, etc.); (4) it would be possible to repeat the survey later. If this survey was selected, on the other hand, the questionnaire created in advance by Statistics Canada could not be modified in light of the day to day realities of the Cree.

The questionnaire prepared by Statistics Canada for the CCHS largely covered the same contents and subject matter as the studies conducted by Santé Québec in 1991. This would allow, among other things, for comparisons between two points in time. There were two main parts to the questionnaire: a common content and an optional content. The common content included those questions which all respondents were asked regardless of the health region where they lived. The optional content consisted of various modules which the provinces could select. In the spring of 2003 there were exchanges between the CBHSSJB and Statistics Canada to identify those modules which might best satisfy the survey's objectives and the need to compare the data with that obtained in the rest of Quebec. For the reader's information, Table 1 lists the optional modules that were chosen for Iiyiyiu Aschii[2].

Table 1
Optional modules of questions chosen by the Cree Board of Health and by the rest of Quebec

Modules	Description	Iiyiyiu Aschii	Quebec
CIH	Changes made to improve health	X	
CMH	Contacts with mental health professionals	X	X
CPG	Problem gambling	X	
DEN	Dental visits	X	X
DIQ	Distress and mental health (Quebec)	X	X
DRV	Driving and safety	X	
FDC	Food choices	X	
FIN	Food insecurity	X	X
HCS	Health care system satisfaction	X	X
HUI	Health utility index	X	X
IDG	Illicit drugs	X	
INC	Income	X	X
MAS	Mastery	X	
NDE	Nicotine dependence	X	X
NUS	Nurses' supplement	X	
PCU	Physical check-up	X	X
PIC	Problems in the community	X	
REP	Injuries	X	
SCA	Smoking cessation aids	X	X
SCH	Smoking – stages of change	X	X
SPV	Spiritual values	X	
STR	Stress	X	
SWA	Satisfaction with availability	X	X

Source: ESCC 2.1, 2003.

2. SAMPLE DESIGN

The sample design was first drawn up jointly by the the Institut national de santé publique du Québec (INSPQ) and the CBHSSJB. It was then submitted to Statistics Canada's experts on survey methods, who reviewed it and adapted it to the particular requirements of the CCHS. The general approach of the sample design was preserved, including the sample size needed for the acquisition of reliable data.

TARGET POPULATION

The component of CCHS Cycle 2.1 involving residents of the Iiyiyiu Aschii health and social service region dealt with a population aged 12 and over residing at home. All native and non-native persons dwelling in the

[2] The complete questionnaire is contained in Appendix A of *Public Use Microdata File Documentation: Canadian Community Health Survey (CCHS) Cycle 2.1,* (Statistics Canada, 2005a).

region's nine communities are included. In 1991 the target population had been exclusively Cree. Persons living in institutions and full-time members of the Canadian Forces were excluded.

SURVEY FRAME

The CBHSSJB provided housing lists for each of the nine Cree communities in the region. These lists were based on three sources: band councils, the school board, and the CBHSSJB. Lists were edited to remove duplications and households which were outside the scope of the survey.

SAMPLE SIZE AND ALLOCATION

A sample of 1,000 individuals was required to arrive at reliable estimates for this health region. Before the distribution of the sample was determined, the nine communities were split into four groups based on their population. The sample of 1,000 individuals was then divided among these groups proportionally to the square root of the estimated population of each one[3]. Group populations were derived from the census of 2001. Then, within groups 3 and 4, the sample size was again distributed according to the square root of the estimated population of each community. Table 2 shows the intended sample distribution and the one which was actually achieved. It must be noted that the initial sample sizes were augmented prior to data collection so as to compensate for the anticipated rate of non-response (~15%) and the number of vacant houses (~5% to 20% depending on the community). The final sample consisted of 920 persons with a distribution among the nine communities that was suitably close to the initially determined proportions.

Table 2
Distribution of the intended and actual sample according to the groups defined in the sample design and the communities

Strata	Communities	Target size (persons)	Actual size (persons)
1	Chisasibi	264	227
2	Mistissini	231	234
3	Waswanipi	90	75
	Waskaganish	106	96
	Wemindji	89	84
	Total	**285**	**255**
4	Whapmagoostui	60	53
	Eastmain	55	58
	Nemiscau	53	43
	Oujé-Bougoumou	52	50
	Total	**220**	**204**
Total Iiyiyiu Aschii		**1 000**	**920**

Source: CCHS 2.1 – Iiyiyiu Aschii, 2003

[3] This procedure maximizes the reliability of the estimates arrived at for each of the four groups and for the nine communities as a whole.

SAMPLING OF HOUSEHOLDS

Within each community, households were selected according to a systematic sampling procedure based on housing lists. The addresses contained in these lists were used to ensure a random spread of selected households in their community. At the household level the sampling fraction varied between 17% and 26% across different communities, the average being around 22%.

SAMPLING OF INTERVIEWEES

Given the size of the desired sample, the total number of households in the region, and the operational constraints, it was agreed that two persons per household should be selected wherever possible. The selection of individuals within a household was designed so as to ensure that all age groups would be well represented in the sample. The selection rule was based on the household's composition. Table 3 describes the rule applied for selecting household members.

Table 3
Selection strategy for individuals to be interviewed according to household composition

Number of persons 12 to 19 years of age	Number of persons ages 20 or over: 0	1	2	3	4	5	6 +
0		A	A	B	B	B	B
1	A	A	C	C	C	C	B
2	A	C	C	C	C	C	B
3	B	C	C	C	C	C	B
4	B	C	C	C	C	C	B
5	B	C	C	C	C	C	B
6 +	B	B	B	B	B	B	B

A: All household members selected (households with 1 or 2 persons).

B: Random selection of two persons aged 12 or over in the household.

C: Random selection of a person aged 12 to 19, and random selection of a person aged 20 or over.

Table 4 shows the expected sample distribution by age group when the selection strategy described in Table 3 is used.

Table 4
Expected sample distribution by age group

AGE GROUP	DISTRIBUTION IN THE 2001 CENSUS (%)	EXPECTED SAMPLE DISTRIBUTION (%)
12-19 YEARS	21.7	21.5
20-29 YEARS	26.5	26.8
30-44 YEARS	28.7	29.9
45-64 YEARS	16.8	15.3
65 YEARS +	6.3	6.6

3. DATA COLLECTION

When the decision to be part of the CCHS was taken in 2003, cycle 2.1 data collection had already been under way throughout Canada since the beginning of the year. Statistics Canada had no objections to CCHS-Cree being involved in this cycle, but speed was essential if certain operational constraints were to be met. These included interviewer training and a relatively brief collection period due to the seasonal nature of the target population's life cycle. August and September were determined as the months when respondent participation would be maximized, since the Crees' fall hunting period begins in the following month.

In a great majority of cases (85%) the questionnaire was administered as a computer-assisted personal interview (CAPI). When it proved impossible to contact the selected individual during the period of data collection in the field, for example due to a prolonged absence, a computer-assisted telephone interview (CATI) was arranged during the last quarter of 2003. The interview took place in the language chosen by the respondent. More than 87% of respondents preferred English, slightly fewer than 7% preferred French, and about 5% preferred Cree.

To learn about other aspects of data collection which were common to all Canadian communities, read Section 6 of the *User Guide for the Public Use Microdata File of CCHS 2.1* (Statistics Canada, 2005a).

4. WEIGHTING[4]

If estimates calculated from the survey data are to be representative of the population concerned–and not merely of the sample–the calculations must include weighting factors, here referred to as survey weights. A survey weight is assigned to each individual in the final sample, i.e. the sample of persons who answered the questionnaire. This weight corresponds to the number of persons in the population as a whole who are represented by the respondent. Table 5 summarizes the various adjustments made according to the weighting strategy in the order in which they were applied.

Table 5
List of weighting adjustments

0	SELECTION WEIGHT, HOUSEHOLDS
1	REMOVAL OF NON-TARGETED UNITS
2	NON-RESPONDING HOUSEHOLDS
3	INDIVIDUAL SELECTION WEIGHT
4	NON-RESPONDING INDIVIDUALS
5	POSTSTRATIFICATION

ADJUSTMENT 0 – SELECTION WEIGHT

Since the initial sample is selected by means of a systematic sampling of households within each one of the nine communities, the selection weight (initial weight) corresponds to the inverse of the probability of a household having been selected. The probability of selection varies from one community to the other. This selection weight is called weight 0.

ADJUSTMENT 1 – REMOVAL OF NON-TARGETED UNITS

Among all the households selected for the sample, a certain proportion was identified in the collection stage as being outside the mandate of the survey. Homes that were either demolished or under construction, and empty, seasonal, or secondary dwellings as well as institutions, are all examples of units not targeted by the CCHS. These units were simply removed from the sample, and only those which were covered by the survey's mandate remained. The latter therefore kept the same weight as in the preceding stage, which is now called weight 1.

ADJUSTMENT 2 – NON-RESPONDING HOUSEHOLDS

It was inevitable that a certain portion of non-responding households would be encountered in the course of data collection. This is normally due to the household refusing to participate, or giving unusable information, or being impossible to contact for the interview. The weights of the non-responding households are redistributed among the respondents according to response groups. Since the only additional information available for the non-responding households is their community, however, the only groups to which they

[4] This section was prepared by Statistics Canada and is an integral part of a working document submitted to the Cree Board of Health and Social Services of James Bay.

correspond are the communities. An adjustment factor is therefore calculated thus within each community:

$$\frac{\sum \text{weight 1 for all households}}{\sum \text{weight 1 for all responding households}}$$

Weight 1 of the responding households is then multiplied by this adjustment factor to produce weight 2. Non-responding households are eliminated from the weighting process as of this point.

ADJUSTMENT 3 – INDIVIDUAL SELECTION WEIGHT

Since the ultimate sampling unit of the CCHS is the individual, the household weight calculated up to this point must be converted into an individual weight. This is achieved by multiplying weight 2 by the inverse of the probability of selection of the individual chosen in the household; the result is weight 3. One should note that two persons per household may be selected for the survey, and that the probability of selection of individuals in the household depends on that household's makeup (see the section on Sampling of interviewees for details about the algorithm governing person selection).

ADJUSTMENT 4 – NON-RESPONDING INDIVIDUALS

As regards the Cree component of the CCHS survey carried out in the Iiyiyiu Aschii communities, an interview may be seen as a two-stage process. The interviewer first obtains a complete list of individuals belonging to the household, and later interviews the person or persons selected out of all the household members. In some cases the interviewers can complete only the first part because the selected person or persons could not be contacted or refused to be interviewed. Such cases are defined as non-responses at the individual level, and an adjustment factor must be applied to the weightings of respondents in order to compensate for this non-response. As in the case of non-response at the household level, the adjustment is applied in the groups defined according to the available characteristics for respondents and non-respondents. All characteristics noted at the stage of listing household members were available for the creation of these groups. The CHAID (Chi-Square Automatic Interaction Detector) algorithm, available in Knowledge Seeker[5], makes it possible to isolate the characteristics which best divide the sample into groups according to the tendency to respond. The following characteristics were used to create the adjustment groups: the number of individuals selected in the household, gender, age group, and community. An adjustment factor was calculated within each group as follows:

$$\frac{\sum \text{weight 3 for all selected persons}}{\sum \text{weight 3 for all responding selected persons}}$$

Weight 3 of responding persons is thus multiplied by this adjustment factor to produce weight 4. Non-responding persons are eliminated from the weighting process as of this point.

ADJUSTMENT 5 – POST-STRATIFICATION

Post-stratification is the last stage required to arrive at the final weighting. This is applied so that the sum of final weights shall correspond to the population estimates for the Iiyiyiu Aschii region for each of the ten target groups with respect to age and gender: i.e. age groups 12-19, 20-29, 30-44, 45-64, 65 and over for males and females. Population estimates used for the reference year 2003 were based on Census counts; this was also the case for birth, death, immigration, and emigration estimates. Thus, weight 4 is adjusted so as to lead to the final weight 5 with the help of an adjustment factor defined as follows:

$$\frac{\text{Population estimate for respondent's age - gender group}}{\sum \text{weight 4 for respondent's age - gender group}}$$

Weight 5 corresponds to the final weighting applied in the survey of residents of the Iiyiyiu Aschii health and social service region. It will be found in the master file with the variable name WTSC_M.

5. DATA QUALITY

RESPONSE RATE[6]

When the units outside the scope of the survey had been removed, a total of 646 households were selected to be part of the CCHS survey carried out among residents of the Iiyiyiu Aschii health region. Out of this number, 581 households agreed to participate; this meant that the response rate at the houshold level was 89.9%. Responding households included 1,074 individuals selected for participation in the survey, of which 920 agreed; at the individual level there was therefore a response rate of 85.7%. Consequently, at the regional level there was a combined response rate of 77.9% in Iiyiyiu Aschii. It should be mentioned that the combined response rate is not obtained by multiplying the household and individual response rates because two persons were selected in some of the households. Table 6 shows the combined response rates for each community as well as the relevant information used to calculate

[5] ANGOSS Software (1995). *Knowledge Seeker IV for Windows - User's Guide.* ANGOSS Software International Limited.

[6] This subsection was prepared by Statistics Canada and is an integral part of a working document submitted to the Cree Board of Health and Social Services of James Bay.

them. We describe below how the various elements of the equation must be handled to determine the correct combined response rates.

HHS = number of households within the scope of the survey

HHR = number of responding households

PS1 = number of persons selected to be part of the survey in households where only one person was selected

PS2 = number of persons selected to be part of the survey in households where two persons were selected

PR1 = number of responders in households where only one person was selected

PR2 = number of responders in households where two persons were selected

RESPONSE RATE AT THE HOUSEHOLD LEVEL

HHRR = HHR/HHS

RESPONSE RATE AT THE INDIVIDUAL LEVEL

PPR = PR1 + PR2

PPS = PS1 + PS2

PPRR = PPR/PPS

COMBINED RESPONSE RATE

CBRR = [PR1 + (PR2/2)]/HHS

Here is a step-by-step example of how the combined response rate was calculated with the information provided in Table 6.

HHRR = HHR/HHS = 581/646 = 89.9%

PPR = PR1 + PR2 = 87 + 833 = 920

PPS = PS1 + PS2 = 88 + 986 = 1,074

PPRR = PPR/PPS = 920/1,074 = 85.7%

CBRR = [PR1 + (PR2/2)]/HHS = [87 + (833/2)]/646 = 77.9%

Table 6
Response rates and number of persons selected for CCHS-Cree with respect to respondents' communities

Communities	Chisasibi	Mistissini	Waskaganish	Wemindji	Waswanipi	Eastmain	Nemiscau	Oujé-Bougoumou	Whapmagoostui	Total Iiyiyiu Aschii
Households										
Target households (Nbr)	175	158	64	56	62	31	30	35	35	646
Responding households (Nbr)	153	144	57	54	50	31	29	30	33	581
Response rates (%)	87.4	91.1	89.1	96.4	80.6	100	96.7	85.7	94.3	89.9
1 person selected										
Persons selected (Nbr)	23	21	10	9	6	1	6	5	7	88
Respondents (Nbr)	23	21	10	9	5	1	6	5	7	87
Response rates (%)	100	100	100	100	83.3	100	100	100	100	98.9
2 persons selected										
Persons selected (Nbr)	260	246	94	90	88	60	46	50	52	986
Respondents (Nbr)	204	213	86	75	70	57	37	45	46	833
Response rates (%)	78.5	86.6	91.5	83.3	79.5	95	80.4	90	88.5	84.5
All persons										
Persons selected (Nbr)	283	267	104	99	94	61	52	55	59	1 074
Respondents (Nbr)	227	234	96	84	75	58	43	50	53	920
Response rates (%)	80.2	87.6	92.3	84.8	79.8	95.1	82.7	90.9	89.8	85.7
Combined response rates (%)	71.4	80.7	82.8	83	64.5	95.2	81.7	78.6	85.7	77.9

Source: CCHS 2.1 – Iiyiyiu Aschii, 2003.

SURVEY ERRORS

Statistics Canada generally identifies two types of error in the interpretation of survey data: errors due to sampling and errors due to causes other than sampling (non sampling errors). For the purposes of the present document we shall deal mainly with errors due to sampling; readers may consult the guide to the microdata files for more complete information on errors due to other causes (Statistics Canada, 2005a). Since the present survey allows estimates to be made from a population sample, its results can be slightly different from those which would be obtained from a complete census of the entire population of Iiyiyiu Aschii. Estimates so produced are subject to sampling error; if we wish to use them to make inferences, we must try to quantify the size of the error. A commonly used method, the coefficient of variation, is the standard error expressed as a function of the estimate with which it is associated. This is the value to be considered when one wishes to know whether the estimate in question is precise enough. How the standard errors of estimates are calculated shall be explained in greater detail in the following section.

6. ACCESS TO DATA AND DETERMINATION OF THE VARIANCE

ACCESS TO DATA

The CBHSSJB has sought to become more autonomous in the matter of health-related data collected in its territory. A special agreement was entered into with Statistics Canada so that the Cree Board of Health would have access to all the data provided by survey respondents. Statistics Canada was nevertheless required to ask each respondent's permission to make his or her information available. In this survey 849 out of the 920 initial respondents, or slightly more than 92%, allowed the statistics agency to share its data with the CBHSSJB. After analysis a data file, called the Share File, was delivered; it is identical to Statistics Canada's master file for the 849 respondents.

The agreement specifies that only the CBHSSJB, its members, and its immediate employees may use these data. Thus the Institut national de santé publique du Québec, which was mandated to perform the data analysis, had no access to personal information. After discussions with representatives of Statistics Canada, the agency agreed that the master file could be queried through its remote access service. The procedure involved relaying SAS programs to Statistics Canada staff, who would then apply them to the master file. After the dissemination rules were checked (see hereunder the section on Dissemination of the results), the results were returned to the users in tabular form. The service greatly enhanced our analytical capabilities, since we were able to use the information provided to the Cree CCHS survey by all respondents.

EVALUATING THE PRECISION OF THE ESTIMATES

Since the sample design is relatively complex, a direct calculation of the precision of the estimates–in this case of the variance–is often fastidious and sometimes impossible. This problem is commonly dealt with by means of resampling techniques such as Jackknife and Bootstrap. For CCHS, Statistics Canada produced along with its master file a set of 500 bootstrap weights which were easy to use through the remote access service. In a nutshell, the procedure is first to calculate the desired estimate 500 times; the variance of these 500 estimates will then be the required variance.

Statistics Canada provides users with SAS or SPSS syntaxes (program Bootvarf_v30.sas) to obtain variances with the bootstrap technique (Statistics Canada, 2005b). We used these programs, with some slight adaptations, to produce the estimates given in the topical publications.

7. DISSEMINATION OF THE RESULTS

DATA PRESENTED IN THE TOPICAL PUBLICATIONS

The data provided in the various tables are proportions. If they are to be considered representative of the population of the Iiyiyiu Aschii health region, all the proportions obtained from the sample must be weighted. The weight variable considered is the one described in the section on Weighting. The proportions in CCHS 2003 for Quebec and Canada as a whole are also weighted for the same reasons.

Moreover, even if comparisons can be made involving two populations with quite different age structures (Iiyiyiu Aschii and the rest of Quebec for example), the proportions presented in all the topical publications have not been adjusted in any way (crude estimates). The estimates provided thus reflect real situations observed in the target populations.

DISSEMINATION GUIDELINES

All estimates in the topical publications are provided in conformity with Statistics Canada's rules governing dissemination. No permission is given to publish when the coefficient of variation exceeds 33.3% or when the number of respondents in a sample used to calculate a numerator is less than 10. In the tables of results the letter "U" (unpublished) is inserted in the place of these imprecise values. As for estimates presenting a high variability, i.e. with a coefficient of variation between

16.6% and 33.3%, their dissemination is restricted. An asterisk (*) marks such estimates, indicating a high variability which must be approached with caution.

Moreover, in order to improve the precision of certain estimates at the sub-regional level (coastal vs. inland communities, sample design strata), small area estimation techniques were considered. Thus the empirical Bayesian composite estimator[7] was used in certain cases; results derived from this estimator are clearly identified in the topical publications.

COMPARISONS WITHIN THE IIYIYIU ASCHII REGION

Descriptive analyses would not be complete without the comparisons made possible by the introduction of different variables. For example, it is highly interesting to study the behaviour of health-related indicators across different age groups or genders. Several proportion tests were carried out for this purpose with the help of the following statistic:

$$Z = \frac{p_i - p_j}{\sqrt{Var(p_i) + Var(p_j)}}$$

where p_i and p_j are respectively the estimated crude proportions of categories *i* and *j* and where $Var(p_i)$ and $Var(p_j)$ correspond to the variances obtained with the bootstrap weightings for the same proportions. It is assumed here that the covariance between the two proportions compared is nil.

The alpha threshold for these tests is 5%. Since Z approximately follows a standard normal distribution, a difference is deemed to be statistically significant if Z is less than -1.96 or greater than 1.96.

In addition to age and gender, tabulation by sub-region or grouping by size of community is often encountered in the topical publications. The communities of Iiyiyiu Aschii have been divided into two sub-regions for purposes of comparison. The coastal communities include Chisasibi, Wemindji, Eastmain, Waskaganish, and Whapmagoostui, while the inland communities include Nemiscau, Mistissini, Oujé-Bougoumou, and Waswanipi. Furthermore, four groups of communities have been defined according to the size of each village population at the time of the survey. These groups correspond to the sample design divisions presented in Table 2: (1) Chisasibi (more than 3,000 inhabitants); (2) Mistissini (2,000 to 3,000 inhabitants); (3) mid-sized communities (1,000 to 2,000 inhabitants), namely Waswanipi, Waskaganish and Wemindji; (4) small communities (fewer than 1,000 inhabitants), namely Whapmagoostui, Eastmain, Nemiscau, and Oujé-Bougoumou. Tabulation involving other variables are also possible and vary in accordance with the topic of each publication.

COMPARISONS BETWEEN THE REGION OF IIYIYIU ASCHII AND THE REST OF QUEBEC

One of the main criteria favouring CCHS was that it made possible a comparison between the Iiyiyiu Aschii region and the rest of Quebec as regards answers given by respondents in both populations. Comparison tests were done with the statistical procedure already described here, but with the following differences taken into consideration.

Age distribution

The population of Iiyiyiu Aschii is much younger than that of the rest of Quebec. Since a number of health indicators are strongly associated with age, it is important to allow for this fact in order to avoid concluding, erroneously, that there exists a difference between the two populations when the observed disparities are entirely due to age. The problem is usually circumvented by adjusting the estimates for the two populations according to the same age distribution. The population of Quebec in 2001 was taken as the reference population, with age groups 12-19, 20-29, 30-44, and 45 and over (data from the Institut de la statistique du Québec). Tests for this type of comparison were made with proportions adjusted according to age.

Data collection

As mentioned in the section on Data collection, the great majority of interviews (85%) in the Iiyiyiu Aschii region were face-to-face encounters. Telephone interviews predominated in the rest of Quebec (St-Pierre & Béland, 2004). Côté and his colleagues (2005) studied the effect of these two types of data collection on the estimates produced, and found large discrepancies for a great number of variables they had analyzed. Thus some major divergences could be due to the difference between data collection methods rather than to a real difference between the populations. To alleviate this difficulty, the authors suggest that the usual alpha threshold of 5% be lowered to 1%.

[7] For more information on this estimator see the working document *Estimations à partir de l'ESCC 2003 Cycle 2.1 - Technique d'estimation sur de petits domaines (Small Area Estimation) - Document de travail : Projet spécial pour la région Iiyiyiu Aschii* translated and reproduced in the Appendix.

The time of collection

The collection period for Iiyiyiu Aschii was largely limited to the two months of August and September, while in the rest of Quebec it extended fairly evenly through the twelve months of 2003. Since some of the survey questions refer to a period prior to the interview[8], the answers can depend greatly on the timing of the interview. For indicators that are related to these questions it is advisable to compare the proportions in Iiyiyiu Aschii with those in the rest of Quebec on the basis of the same period, that is to say August and September.

In conclusion, comparison tests between the Iiyiyiu Aschii region and the rest of Quebec for the 2003 CCHS were carried out with an alpha threshold of 1% on proportions adjusted according to age. As for indicators which can be affected by the time when interviews occurred, the proportions compared are calculated solely from interviews done in August and September.

HOW THE SITUATION HAS CHANGED SINCE 1991: SANTÉ QUÉBEC SURVEY (1991) AND CCHS-CREE (2003) COMPARED

A primordial goal is to analyze how the health of Iiyiyiu Aschii residents has evolved with time. In relatively recent times the 1991 Santé Québec survey of the James Bay Cree was the only large-scale study to have covered several aspects of health and its determining factors. It therefore constitutes an undeniable benchmark. The decision to go ahead with CCHS in 2003 lowered the potential for comparative analyses, however, as the topics covered by the two surveys were not all identical. Moreover, questions can always be raised about the suitability of these surveys for comparison, since the wording of questions and the choices offered for responses can vary. Each indicator must therefore be treated separately to verify that the comparison is valid. It would be overly ambitious to review here all the indicators used for the purpose of time-based comparisons; consequently, readers are referred to the various topical publications in which these indicators are analyzed. On the other hand, it is important to mention the major guidelines observed when these comparisons are undertaken.

Target populations

The two surveys did not cover exactly the same populations. While CCHS-Cree 2003 deals with all persons aged 12 and over (Cree and non-Cree) living in private households, the 1991 Santé Québec survey covered only Cree aged 15 and over living in private households. To make comparison possible, the proportions from CCHS-Cree 2003 must be calculated strictly for Cree persons aged 15 and over. Since a question on ethnic identity was included, it is possible to distinguish between Cree and non-Cree respondents in CCHS.

Age distribution

Because the age structure of the Cree population has changed over the last twelve years it is preferable, for the same reasons as given in the section on the Dissemination of the results, to work with proportions adjusted according to age to ensure the validity of the results. The reference population remains that of Quebec in 2001, but the age groups must be established as follows: 15-24, 25-44, 45-64, 65 and over.

Questionnaires

Unlike CCHS 2003 which had only one questionnaire, the 1991 Santé Québec survey incorporated several collection methods, the principal ones being two questionnaires: the first one, administered in a face-to-face situation with a single person answering for all members of a household, and the second one, theoretically self-administered but mainly done in face-to-face fashion with each individual 15 years old or over in a selected household. The collection method–face-to-face interviewing–was thus the same for the two surveys and the usual alpha threshold of 5% for tests was therefore maintained.

To sum up, comparison tests between the 2003 survey (CCHS-Cree) and that of 1991 (Santé Québec) were done with a 5% threshold on proportions adjusted with respect to age groups and calculated exclusively for Cree persons aged 15 and over.

REFERENCES

Côté L., Courtemanche, R. & B. Caron (2005). *Comparabilité entre les cycles 1.1 et 2.1 de l'Enquête sur la santé dans les collectivités canadiennes : Impact du changement apporté à la répartition de l'échantillon selon la base de sondage*. Quebec : Institut de la statistique du Québec.

Santé Québec. Daveluy, C., Lavallé, C., Clarkson, M., & Robinson, E. (dir.) (1994). *A Health Profile of the Cree, Report of the Santé Québec Health Survey of the James Bay Cree 1991*. Montreal : Ministère de la Santé et des Services sociaux, Government of Quebec.

[8] Regarding physical activity, for example, respondents were asked if they had engaged in one during the three previous months.

Statistics Canada. (2003). *Canadian Community Health Survey (CCHS), Cycle 2.1*. Ottawa: Health Statistics Division. [Online]. http://www.statcan.ca/english/ concepts/health /cycle2_1/index.htm

Statistics Canada (2005a). *Public Use Microdata File Documentation - Canadian Community Health Survey (CCHS) Cycle 2.1* (2003). Statistics Canada.

Statistics Canada (2005b). *Estimation of the Variance Using Bootstrap Weights: User's Guide for the BOOTVARF_V30.SAS program (version 3.0)*. Statistics Canada.

St-Pierre, M. & Y. Béland (2004). *Mode effects in the Canadian Community Health Survey: a comparison of CAPI and CATI. 2004 Proceedings of the American Statistical Association Meeting, Survey Research Methods Section.* Toronto: American Statistical Association.

APPENDIX

Estimates based on CCHS 2003 Cycle 2.1

Small area estimation:
Special project for Iiyiyiu Aschii region

From: *Estimations à partir de l'ESCC 2003 Cycle 2.1 - Technique d'estimation sur de petits domaines (Small Area Estimation) - Document de travail : Projet spécial pour la région* Iiyiyiu Aschii

Context

The Iiyiyiu Aschii health region, a remote region in northern Quebec, is included for the first time in the Canadian Community Health Survey (CCHS) Cycle 2.1. The Cree Board of Health and Social Services of James Bay (CBHSSJB) purchased a supplementary sample which made it possible to produce reliable estimates on a number of health-related variables or indicators for the region as a whole. The choice of question modules closely followed that of Quebec, with some exceptions made in the light of the region's particular situation. In most cases, therefore, it is possible to make comparisons with Quebec.

Since the sample design covers each of the nine communities in Iiyiyiu Aschii, we wish to go further by producing estimates on small domains involving sub-regions (coastal vs. inland communities), groups of communities based on size, and even single communities. While the usual estimation technique derived from the sample design, here called a direct estimator, can be applied in certain cases, it often leads to a high variability due to the restricted sample size in these small areas. In most cases the values of the coefficients of variation do not meet Statistics Canada's guidelines (see the CCHS 2.1 User Guide for the Public Use Microdata File (Statistics Canada, 2005a)). However, the estimates can be made more precise by the use of a composite estimator described in the present document.

Small domains

Before describing the estimation technique properly, it is essential to define the small domains that are of most interest to us here. These areas correspond to subdivisions of the sample determined according to groups of communities and categories of age and gender. More exactly, we wish to produce estimates for age-gender combinations as well as for the following groups of villages:

Sub-region:

- Coastal communities:
  Chisasibi, Wemindji, Eastmain, Waskaganish, Whapmagoostui.
- Inland communities:
  Nemiscau, Mistissini, Oujé-Bougoumou, Waswanipi.

Groupings according to community size:

1. Chisasibi (more than 3,000 inhabitants);
2. Mistissini (2,000 to 3,000 inhabitants);
3. Mid-sized communities (1,000 to 2,000 inhabitants), i.e. Waswanipi, Waskaganish, and Wemindji;
4. Smaller communities (fewer than 1,000 inhabitants), i.e. Whapmagoostui, Eastmain, Nemiscau, and Oujé-Bougoumou.

Similarly for the eight age-gender combinations (12-19; 20-29; 30-44; 45 and over).

Description of the method applied for small area estimation: the empirical Bayesian composite estimator

The method described hereunder is based mainly on the work of Chattopadhyay, Lahiri, Larsen, and Reimnitz (1999)[9]. The same method was used by Statistics Canada to produce low proportion estimates concerning four sub-regions of Prince Edward Island for Cycle 1.1 of CCHS. With this method it is possible to make a noticeable improvement in the precision of estimates while minimizing bias.

Notation:

i : subscript identifying the sub-region

k subscript for the age-gender combination

j : subscript for a respondent ($j = 1,...,n_{ik}$) where n_{ik} is the sample size of sub-region i for the age-gender combination k

y_{ikj} : jth observation (0 or 1) of the kth age-gender combination in sub-region i

[9] Chattopadhyay, M., P. Lahiri, M. Larsen, and J. Reimnitz (1999). *Composite Estimation of Drug Prevalences for Sub-State Areas. Survey Methodology,* 25, no. 1, pp. 81-86.

w_{ikj} : weight associated with individual of observation y_{ikj}

Using a linear combination of direct estimators for the region and sub-region, one can define the composite estimator of the *k*th age-gender combination of sub-region *i* in the following manner:

$$\hat{\pi}_{ik}^{com} = \lambda_{ik}\hat{\pi}_{ik}^{dir} + (1-\lambda_{ik})\hat{\pi}_{k}^{dir}$$

where $\hat{\pi}_{ik}^{dir}$ represents the direct estimator of the *k*th age-gender combination of sub-region *i* and $\hat{\pi}_{k}^{dir}$ represents the direct estimator of the *k*th age-gender combination of the region with a coefficient λ_{ik} between 0 and 1.

The following model makes it possible to formulate an empirical Bayesian estimator of $\hat{\pi}_{ik}^{com}$ and thus to determine the value of constant λ_{ik} .

Model

- Depending on π_{ik} , observations y_{ikj} are uncorrelated with one another with $\mathrm{E}(y_{ikj}/\pi_{ik}) = \pi_{ik}$ and $\mathrm{Var}(y_{ikj}/\pi_{ik}) = \pi_{ik}(1-\pi_{ik})$
- The π_{ik} are similarly non-correlated with as expectation $\mathrm{E}(\pi_{ik}) = \mu_k$ and as variance $\mathrm{Var}(\pi_{ik}) = d\mu_k^2$
- π_{ik} ~ Uniform $(\mu_k - \alpha\mu_k, \mu_k + \alpha\mu_k)$ where the value of α is between 0 and 1.

  Thus $d = \alpha^2/3$

In other words, the model suggests that the real proportions of the sub-regions (for each age-gender combination) observe a uniform law in the interval $(\mu_k - \alpha\mu_k, \mu_k + \alpha\mu_k)$, and thus centred around the real proportion of the region (μ_k). The value α is determined with data obtained from the other Quebec's regions in CCHS Cycle 2.1. This will be discussed in greater detail in the next section.

The empirical Bayesian estimator of $\hat{\pi}_{ik}^{com}$, under the abovementioned model and the squared error loss function, is given by

$$\hat{\pi}_{ik}^{eb} = \hat{\lambda}_{ik}\hat{\pi}_{ik}^{dir} + (1-\hat{\lambda}_{ik})\hat{\pi}_{k}^{dir}$$

where $\hat{\lambda}_{ik} = \dfrac{d(\hat{\pi}_k^{dir})^2}{\left[d(\hat{\pi}_k^{dir})^2 + c_{ik}\left(\hat{\pi}_k^{dir} - (d+1)(\hat{\pi}_k^{dir})^2\right)\right]}$

with $c_{ik} = \dfrac{\sum_{j=1}^{n_{ik}} w_{ikj}^2}{\left(\sum_{j=1}^{n_{ik}} w_{ikj}\right)^2}$ and $d = \alpha^2/3$

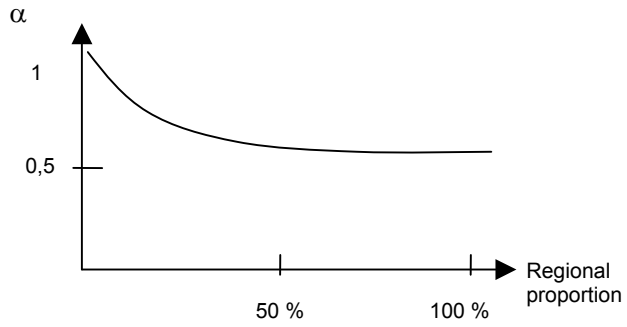
The choice of α

In the last formula described above, all terms can be easily obtained except for constant α. This represents, according to the model, the uniform distribution around the regional proportion for a given characteristic. It can easily be deduced that this value will depend directly on the regional proportion. Indeed, for a similar range, the greater is the proportion, the smaller the value α will be. To verify this statement and determine what α values should be applied for given regional proportions, we can use the data of CCHS 2.1 to estimate regional variations. Three regions in Quebec–those of the Capitale-Nationale, the Outaouais, and the Laurentides–have purchased supplementary samples in order to produce estimates at sub-regional levels, namely the CLSC districts. We can therefore evaluate the variability of the estimates for the areas composed of the CLSC and the eight age-gender combinations, as well as for various characteristics, around the estimated values for the corresponding region. The characteristics involved are:

- Perception of state of health: excellent and very good.
- Obesity: BMI > 30
- Smoking
- Physical leisure activity: inactive
- Education level: less than a secondary diploma

For example, we calculated the variability among the proportions of smokers of each of the eight CLSCs in the Capitale-Nationale region around the regional proportion of smokers among males from 12 to 19 years of age. By repeating this exercise for the other age-gender combinations, the various characteristics singled out, and the three target regions, we end up with several intra-

regional measures of variability for a given scale of regional proportion. We can then determine the value α for a given proportion or regional proportion by taking the 95th percentile of these measures and dividing by this proportion. Putting the α values so obtained as a function of their associated regional proportions, we get a graph of this form:



In view of the results obtained, we suggest estimating the value α for a given proportion with the help of a quadratic regression model; the equation is given below. We must add that the characteristics giving a similar proportion will have the same α value.

$$\alpha = 1.0689 - 0{,}03073\hat{\pi}_k^{dir} + 0.00028(\hat{\pi}_k^{dir})^2 \quad ; R^2 = 0.6758$$

Since α has a maximum of 1, any value greater than 1 produced by this equation shall be reduced to 1.

Variance estimation

Statistics Canada suggests the use of bootstrap weights to estimate the variance of the CCHS direct estimators. To obtain an estimate of the variance of the composite estimator, the same bootstrap weights can readily be used with a slight adaptation. One must calculate the composite estimator separately for each bootstrap sample (that is to say 500 times), where b represents the bootstrap sample, b = 1 to 500.

$$\hat{\pi}_{ik(b)}^{eb} = \hat{\lambda}_{ik(b)}\hat{\pi}_{ik(b)}^{dir} + \left(1 - \hat{\lambda}_{ik(b)}\right)\hat{\pi}_{k(b)}^{dir}$$

The variability observed among these 500 estimates will then be the desired variance estimator.

Dissemination rules adapted for composite estimators

Statistics Canada imposes dissemination rules for any proportions obtained from the CCHS 2003 file by the direct method: the coefficient of variation (CV) of the estimate must be less than 33% and the sample size for the estimate must be at least 30 respondents (10 if bootstrap weights are used). We shall try whenever possible to produce and present direct estimates in accordance with these rules for small areas. On the other hand, if and only if these requirements cannot be met, we shall proceed with the production of composite estimates as described above. These estimates must nevertheless satisfy minimal standards of precision, i.e. they must have a CV under 33% and the minimum size for each table cell must be 10. Cases may arise where there are empty cells even if a small area estimation method is used. In most cases, composite estimation methods will ensure that the CVs will be sensibly reduced (with improved precision) and so allow the dissemination of estimates which otherwise (with direct estimation) could not be presented. The proportions obtained through the composite estimation method must be clearly marked in the tables and an appropriate note attached thereto.