Methodology

SURVEY METHODOLOGY FOR THE 2003 AHS III

The Adolescent Health Survey, first conducted in 1992, was devised to identify critical issues facing BC youth, and to contribute to efforts to promote better health among this age group. The second cycle of the survey (AHS II) was conducted in 1998, to update and expand on the provincial baseline data started with the first cycle. AHS III, conducted in 2003, is the third cycle of the survey. Its content is similar to that of the previous cycles in order to allow the tracking of trends amongst BC students. As well, the content was expanded, somewhat, to reveal emerging health issues.

Target Population and Survey Design

As in the past two cycles, the initial target population for AHS III included all students who were enrolled in grades 7 through 12 (in both public and independent schools) this time during the 2002/03 school year. The very low response rate amongst independent schools forced their exclusion from the final AHS III sample. Accordingly, the population represented by the AHS III sample is all 290,000 students enrolled in grades 7 through 12 in public schools during the 2002/03 school year. Readers are referred to the section "Comparability of AHS III to Previous Cycles" for further explanation.

The sample design for AHS III was similar in size and scope to that used for AHS II. The sample frame was, in essence, a list of all classrooms in the province, stratified by geography and by grade (7 through 12). Independent random samples of classrooms were selected in each stratum. All students enrolled in each selected classroom, fell into the sample.

The sampling frame used for AHS III was the BC Ministry of Education list of all schools, which included enrolment counts by grade for the 02/03 school year. The AHS III was required to produce statistically reliable estimates at each grade level, for each of 16 Health Service Delivery (HSD) areas. Accordingly, the frame was stratified geographically by these regions, and then by grade. Sample sizes were calculated for each of these strata to ensure that the resulting regional estimates by grade would have maximum standard errors of 3.5%. Because of these very specific geographic requirements, the sampling rate was not constant for all areas of the province. It varied from 1 in 3 students for regions with small student populations, to 1 in 8 students for regions with large populations.

Within a HSD area, the required sample was apportioned to each school district, in proportion to the district's enrolment. The sampling of classrooms from a district's schools was roughly proportional to the schools' enrolments. This meant that within a HSD area, all students had roughly the same probability of selection, regardless of their grade level, the size of the school they attend, or the size of the enrolment in their particular school district. Of the 59 school districts in the province, 45 permitted access to the selected schools and classrooms. These 45 school districts contain 72% of all students enrolled in grades 7 through 12 in public schools in the province. In order to maintain the sample size required for each HSD area, classrooms selected for non-participating districts were reassigned to other randomly selected schools from participating districts in the same HSD area (maintaining the equal probability of selection for students). This sample re-allocation technique was used in all areas of the province with the exception of the Fraser Valley, where the relatively high rate of school district non-participation made it impractical. In total, 40,040 students from 1,557 different classrooms, in 440 different schools, fell into the AHS III sample.

Although the AHS III sample was designed with the current boundaries used by the BC Ministry of Health, the basic geographic building block is the school district. The aggregation of school districts allows estimates to be produced for various geographies, including the eight AHS provincial areas used for previous AHS cycles.

Survey Administration and Data Collection

Data collection for AHS III was similar to that used for AHS I and AHS II. Data collection occurred in school, during the March to June 2003 period. Public health nurses were responsible for the data collection in their region.

Parents of students in classes selected for the survey were notified in advance, via letters, concerning the conduct of the survey and the voluntary nature of student participation. Each parent contacted also had an opportunity to review the survey questionnaire by visiting the school office. If parents did not wish their son/daughter to participate they could indicate so on the consent forms supplied. At the discretion of the school district, parents provided either active or passive consent. Active consent meant that sampled students were permitted to

complete a survey questionnaire only if they had received written approval from their parents. Passive consent meant that sampled students were excluded from participating in the survey only if their parents had indicated that they did not wish their son/daughter to participate. All surveying was completely voluntary, anonymous and confidential. Student names were not recorded and students did not put their names on the questionnaires.

The survey was administered in school classrooms during regular school hours. A health unit representative or other trained survey administrator, was on hand to provide instructions for completing the questionnaire, answer student questions, and ensure response privacy. Survey administrators were given specific instructions on how to administer the survey. These instructions and the answers/assistance given to students' queries were standardized across administrations in all grades and schools. In addition, all administrators were provided with forms and detailed instructions for collecting information on classroom enrolment, absenteeism, and parent/student refusals for use in calculating response rates and weighting the survey data.

Response Rates

For school surveys such as this, the target sample consists of all the students enrolled in the 1,557 classrooms that were randomly selected from among all of the classrooms in the participating school districts. The response rate represents the percentage of students enrolled in the selected classes that completed the survey and provided useable information. To the extent that the response rate is less than 100%, the difference is usually accounted for by the proportion of enrolled students who did not respond to the survey for one of the following reasons: a) they were absent on the day the survey was given; b) their parents either failed to provide a consent form or refused their consent; or c) the students themselves refused to take part in the survey.

The overall response rate for the province was 76%. In school districts where school board officials required active parental consent, the

resulting response rate was just 58%. On the other hand, in school districts that permitted passive consent, the response rate was 84%. This difference suggests that a significant portion of non-response to the survey occurred because parents failed to return consent forms, rather than actual refusals by parents. In fact, 7% of all students sampled for the survey failed to return active consent forms to the school prior to surveying, whereas only 2% of parents refused their consent. Additional sources of non-response are attributable to absenteeism (12%) and refusals by students (1%). Finally, a relatively small number of students (< 1%) provided partial or inadequate data records because they either did not answer a significant number of questions or failed to give honest and trustworthy answers. Since absenteeism is the largest source of non-response, any differences between students present or absent on survey administration day may have an impact on the results.
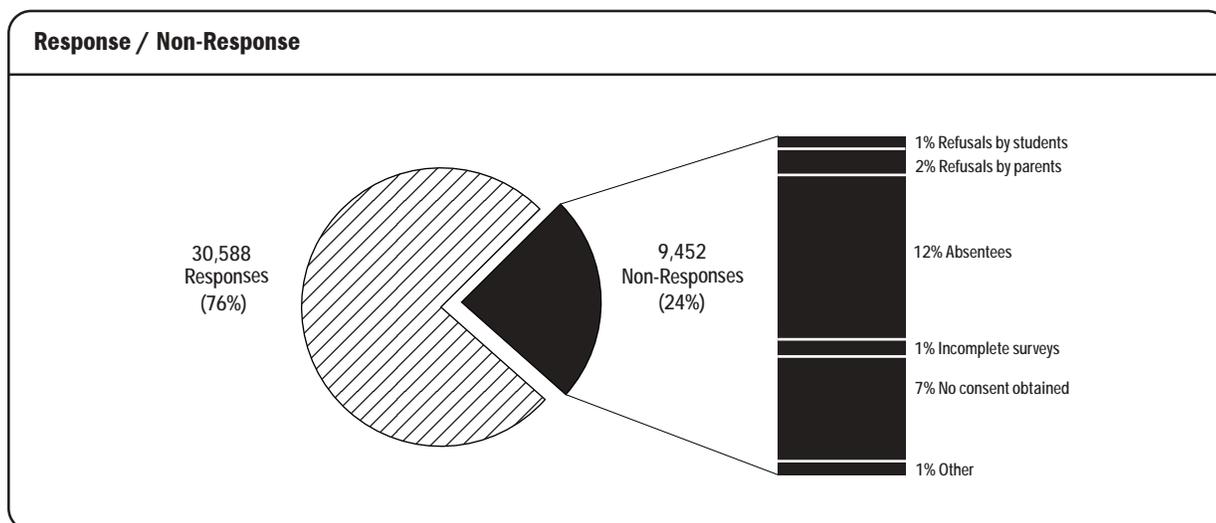
There was virtually no variation in response rates by grade, indicating that the target population's grade structure is well represented by the AHS III sample.

Health Service Delivery Areas (16 provincial regions):
Response rates varied slightly amongst HSD areas (62% - 86%), according to whether a region's school districts had adopted active or passive consent. The only exception of note occurred in the Northern Vancouver Island HSD area, where a high absentee rate coupled with a high incidence of consent forms not returned, forced the response rate below 50%. Any regional differences in response rates were accounted for in the weighting.

Geographic Areas (8 provincial regions):
For the 8 provincial geographic areas, response rates varied slightly from area to area (60% - 84%), according to whether a region's school districts had adopted active or passive consent. Any regional differences in response rates were accounted for in the weighting.

Response / Non-Response

30,588 Responses (76%)

9,452 Non-Responses (24%)

1% Refusals by students
2% Refusals by parents
12% Absentees
1% Incomplete surveys
7% No consent obtained
1% Other

2003 AHS III Provincial Response Rates

	2002 - 03 Enrollment	# of Students Solicited	# Absent	# of Refusals by Parents	# of Refusals by Students	# of Active Forms Not Returned	# of Other†	# of Incomplete or Unuseable Questionnaires	# of Usable Questionnaires	Response Rate
Participating School Districts (45 SD)	207222	40040	4654	865	460	2802	298	373	30588	76%
All Active Consent (12/59 SD)	37049	9825	1189	255	223	2265	155	71	5667	58%
All Passive Consent (24/59 SD)	118515	23155	2666	449	171	0	94	257	19518	84%
Both Active & Passive Consent (9/59 SD)	51658	7060	799	161	66	537	49	45	5403	77%
Non-Participating School Districts (14 SD)	79665	0								0%
All School Districts	286887	40040	4654	865	460	2802	298	373	30588	76%

† Other = reason for not completing survey is unknown, student had unique educational needs (ie. ESL, special needs) and/or survey administrator error in providing response information

2003 AHS III Response Rates for Health Areas, Geographic Areas and Grade Levels

	2002 - 03 Enrollment	# of Students Solicited	# of Usable Questionnaires	Response Rate		2002 - 03 Enrollment	# of Students Solicited	# of Usable Questionnaires	Response Rate
Ministry of Health Areas (16 Provincial HSD Areas)					**Grade Summary**				
Northwest (3/6 SD)	7142	1897	1252	66%	Grade 7	46267	8010	5933	74%
Northeast (0/3 SD)	5509	0	0	0%	Grade 8	47154	6575	5153	78%
Northern Interior (3/3 SD)	12641	3762	2317	62%	Grade 9	47819	6627	5217	79%
Thompson-Cariboo (6/6 SD) *	16734	3422	2588	76%	Grade 10	47510	6063	4782	79%
Okanagan (6/6 SD) *	23332	3744	2614	70%	Grade 11	50176	6534	4791	73%
Kootenay Boundary (4/4 SD) **	6050	2057	1699	83%	Grade 12	47961	6231	4712	76%
East Kootenay (3/3 SD) **	6392	2082	1774	85%	All Grades	286887	40040	30588	76%
North Vancouver Island (3/4 SD)	9436	1848	811	44%	**Geographic Areas (Used in AHS I, II, & III)**				
Central Vancouver Island (4/4 SD)	17701	3607	2953	82%	Greater Vancouver	137108	13077	10927	84%
South Vancouver Island (4/4 SD)	20113	2730	2300	84%	Capital	20113	2730	2300	84%
North Shore / Coast Garibaldi (5/6 SD)	18740	3790	3115	82%	Fraser Valley	18868	345	220	64%
Vancouver (1/1 SD)	27407	3605	3096	86%	Interior	36474	6132	4625	75%
Richmond (1/1 SD)	12201	3357	2883	86%	Kootenay	12442	4139	3473	84%
Simon Fraser (3/4 SD)	38529	2970	2299	77%	Upper Island	32998	6924	4897	71%
South Fraser (1/3 SD)	46092	824	667	81%	Northwest	7142	1897	1252	66%
Fraser Valley (1/4 SD)	18868	345	220	64%	Northeast	21742	4796	2894	60%
All Regions	286887	40040	30588	76%	**All Geographic Areas**	**286887**	**40040**	**30588**	**76%**

* SD #58 and #83 both span two health areas ** SD #8 spans two health areas

Sample Representation

With respect to sample representation, the potential bias introduced by the non-participating school districts, is of greater concern than response rates. The 14 school districts that refused access to their schools represent 28% of the student population enrolled in grades 7 to 12 in public schools. In essence, students in these districts were not given the opportunity to participate, and, hence they are not represented at all in the AHS III. To the extent that students from these districts possess attributes that differ from those of students from the 45 participating school districts, their lack of representation will have an impact on the accuracy of population estimates derived from the AHS III. The bias introduced by this under-coverage of the population can be speculated on but never adequately measured.

The non-participation of school districts exhibited geographic clustering, which may affect regional estimates more profoundly than provincial estimates. The "coverage rate" is the proportion of the target population covered by participating school districts and, hence, actually represented by the sample.

Health Service Delivery Area (16 provincial regions):
Coverage rates were at or near 100% in all HSD areas, with the exception of the Northeast HSD area and those comprising the Fraser Valley. The three school districts in the Northeast HSD area did not participate, so separate results are not available for this area. The 57% coverage in the Simon Fraser HSD area is considered adequate, but the very low coverage rate in the South Fraser and Fraser Valley HSD areas meant that these region's estimates had to be suppressed. The Fraser Valley and South Fraser respondents are, however, included in the provincial estimates and the 5 provincial Health Authority/Ministry of Children and Family Development Regions' estimates, in their correct proportions.

Geographic Areas (8 provincial regions):
Coverage rates were well above the provincial average of 72% in all geographic areas of the province with the exception of the Greater Vancouver area and the Fraser Valley. The 61% coverage in Greater Vancouver is considered adequate, but the low coverage rate in the Fraser Valley meant that this region's estimates had to be suppressed. The Fraser Valley respondents are, however, included in the provincial estimates, in their correct proportions.

Data Processing and Weighting

Some of the students' data records were deleted from the final sample during the compilation or data processing stage. This was done to maximize the validity and accuracy of the survey estimates. Approximately 1% of the survey records were deleted based on one or more of the following criteria:

- respondents who failed to answer the majority of questions;
- respondents who provided a number of inconsistent or contradictory answers.

Survey data records from the 30,588 students that responded to the AHS III have been adjusted (weighted), so that they provide an accurate representation of all 290,000 public school students in grades 7 through 12 in all regions of the province–including those attending schools in non-participating school districts–as well as students in schools and classrooms who were not among those randomly selected from the participating districts. The weight attached to each respondent ensures his or her appropriate representation, and can be thought of as the number of students in the population that the respondent represents. The weight is the product of three broad factors:

- the probability of selection (ie, being in a classroom randomly selected for inclusion);
- the non-response adjustment;
- population readjustments, used to ensure population representation according to provincial region, school district, and grade level.

Just as the sampling rates and school district coverage rates vary geographically, so do the weights. As a consequence, accurate results can only be produced from the AHS III sample, if the weights are used in tabulations.

(During processing, it was noticed that the AHS III sample exhibited a 1%-2% shift in the gender distribution of students, when compared to BC Ministry of Education enrolment data. As a test, a raking ratio adjustment was made to the weights, to correct this slight imbalance. This weighting adjustment made no appreciable difference to the estimates so it was left off.)

Missing Data

As in all surveys, respondents occasionally refuse, or are just unable to provide a valid response to every question. Examples include "don't know", "not sure" or "does not apply" responses, and instances of not providing any response or answer at all (ie, leaving it blank). Levels of this type of non-response to individual questions on the AHS III are very low in general–about 1 to 2% of the sample. Non-responses are not shown in reports and fact sheets from the survey; instead they are allocated (imputed) to the other answer categories proportionally. But where "don't know" or "not sure" is an informative response, such as with knowledge questions, this response option is routinely reported.

Release Criteria and Significance Testing

Survey estimates from the AHS III, in the form of proportions or percentages, are based on a sample of students that were randomly selected. Somewhat different figures might have been obtained if a complete census of the target population had been conducted. The difference between an estimate based on a sample and the value obtained from a census taken under similar conditions (ie, same questionnaires, administration procedures and data processing methods) is called the sampling error. An indication of the magnitude of the sampling error, or the relative precision of an estimate, can be approximated from the sample itself using a measure called the standard error (SE).

The complex design of the AHS III sample makes it inappropriate to calculate standard errors based on simple random sampling theory.

The standard errors for AHS III have been adjusted by incorporating a factor called the "design effect" for the survey. The design effect reflects the increase in the sampling error (relative to a simple random sample of the same size) caused by using different sampling rates in various geographies, and by sampling entire classrooms rather than individual students. The design effects ranged from a high of 2.75 for provincial estimates to 1.5 for most smaller geographies.

The publication or release of estimates from the AHS III is governed by the size of the standard errors associated with them:

- survey estimates with SE's less than 4.99 are published without qualification;

- survey estimates with SE's between 5.00 and 12.49 are published with reservation because of potentially high sampling variability (indicated by a * to denote a qualified estimate that should be used with caution);

- survey estimates with SE's exceeding 12.5 are suppressed (indicated by a #) to denote that they are not releasable.

Differences in the size of proportions or percentages from the survey were tested for statistical significance, before they are noted in the text of reports or fact sheets. These include comparisons of subgroups within AHS III (eg. comparisons by gender or grade for AHS III), as well as comparisons between AHS III and previous cycles (AHS I and AHS II). The difference between two estimated proportions (P1 - P2) is considered to be statistically significant at the 5% level (p<.05) if the difference in the size of the two estimates divided by the sum of their combined standard errors exceeds the t-alpha value of 1.965. The following equation was used for this purpose:

$$\frac{|P1 - P2|}{\sqrt{(SE_{P1}^2 + SE_{P2}^2)}} > 1.965$$

It is important to note that, because of the large sample size obtained in the AHS III, many of the differences observed may well prove to be statistically significant, but have no programmatic or policy importance. Therefore, among the various reports and fact sheets produced from the AHS III, not all statistically significant differences are noted; however, all differences that are noted in the text have been tested for statistical significance.

Finally, all survey estimates published from the AHS III have been rounded to whole numbers, as this aids readability and provides an appropriate level of precision. Due to rounding, however, the percentages shown in some figures and tables may not always sum to exactly 100%.

Comparability of AHS III to Previous Cycles

From a methodological perspective, there are two aspects of the AHS III where differences from previous cycles could affect comparisons:

- the exclusion of students from independent schools from the target population of AHS III;

- differences in the student population actually covered by the participating school districts.

Small studies were undertaken to determine whether either factor had a large enough impact on the estimates to warrant a caution about comparability.

The impact of excluding students from independent schools, was approximated by producing estimates from the 98 AHS II sample, both with and without independent school students. This study produced two important findings:

- firstly, independent school students were probably under-represented in the 98 AHS II (they accounted for less than 2% of the student population);

- secondly, because they represented such a small group, their exclusion had virtually no impact on the AHS II estimates.

In summary, the conceptual change in the target population of the AHS III to include only public school students, does not appreciably affect its comparability to previous cycles.

Coverage of the student population in grades 7 through 12, improved steadily (albeit, slightly) throughout the three cycles of the AHS. This improvement raises the concern about whether the 'covered population' is, essentially the same for the three cycles, and, hence, whether trends can be tracked with accuracy. In the AHS III, participating school districts accounted for 72% of the overall BC student population. Comparisons with previous cycles showed that, of the 28% under-covered, a full 17% was accounted for by persistent non-participation (that is, school districts that had refused for all three cycles). While the high rate of persistent non-participation causes concern about the accuracy of any one cycle, it provides some assurance that the population covered by the three cycles is essentially the same. Notwithstanding this, the accurate tracking of trends could be affected, if students in persistently non-participating school districts exhibited changes over time that were vastly different from the rest of the province's students.

Prepared by Rita Green, Survey Consultant, Statistics Canada

The McCreary Centre Society
3552 E. Hastings Street
Vancouver, BC V5K 2A7
tel (604) 291-1996
fax (604) 291-7308
mccreary@mcs.bc.ca
www.mcs.bc.ca