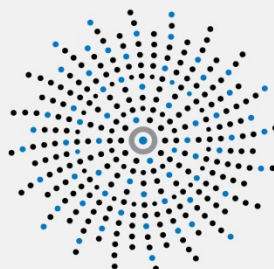# MICRODATA USER GUIDE

2018/2019 Canadian Student Tobacco, Alcohol and Drugs Survey

September 2019

PROPEL
CENTRE FOR
POPULATION
HEALTH IMPACT

Suggested Citation

Burkhalter, R., Thompson-Haile, A., Cooke, M. (2019). 2018/2019 Canadian Student Tobacco, Alcohol and Drugs Survey Microdata User Guide. Waterloo, Ontario: Propel Centre for Population Health Impact, University of Waterloo, 1-39.

Suggested Acknowledgement when using CSTADS Data

> *Data used for this research were taken from Health Canada's Canadian Student Tobacco, Alcohol and Drugs Survey (CSTADS), which was conducted for Health Canada by the Propel Centre for Population Health Impact at the University of Waterloo. Health Canada has not reviewed, approved, nor endorsed this research. Any views expressed or conclusions drawn herein do not necessarily represent those of Health Canada.*

PLEASE NOTE: For information purposes only, Health Canada would appreciate receiving advanced copies of planned publications arising from CSTADS data at least 3 weeks prior to the publication date. Send a copy by mail or email to the Technical Authority.

For further information regarding this project, please contact:

Martin Cooke (CSTADS Principal Investigator)
Propel Centre for Population Health Impact University of Waterloo
200 University Avenue West Waterloo, ON N2L 3G1
519-888-4567 x36585
cooke@uwaterloo.ca
www.cstads.ca

Please direct any questions about the data set or its use to:

University of Waterloo	*Health Canada*
Robin Burkhalter (CSTADS Analyst)	Rachel Bennett (CSTADS Technical Authority)
Data Analyst	Senior Epidemiologist
200 University Avenue West	Tobacco Control Directorate 150 Tunney's
Waterloo, Ontario N2L 3G1	Pasture Driveway Ottawa, ON K1A 0K9
(519) 888-4567 ext. 36632	(613) 330-8654
rjburkha@uwaterloo.ca	hc.cstads.questions-ectade.sc@canada.ca

PLEASE BECOME FAMILIAR WITH THE CONTENTS OF THIS DOCUMENT BEFORE PUBLISHING OR OTHERWISE RELEASING ANY ESTIMATES DERIVED FROM THE 2018/2019 CSTADS PUBLIC USE MICRODATA FILE.

Table of Contents

List of Tables

List of Appendices

1.0 Introduction

Health Canada's 2018/2019 Canadian Student Tobacco, Alcohol and Drugs Survey (CSTADS; formerly the Youth Smoking Survey) is a Canadian, biennial, school-based survey of a generalizable sample of students in grades 7 through 12. CSTADS (YSS) has been coordinated centrally by the Propel Centre for Population Health Impact (Propel) at the University of Waterloo in partnership with Health Canada since 2004. The 2018/2019 CSTADS was implemented in schools between October 2018 and June 2019 by provincial teams located in the participating provinces, with support and leadership from the following CSTADS co- investigators:

Dr. Linda Rohr	Memorial University of Newfoundland
Dr. Donna Murnaghan	Thompson Rivers University
Dr. Jo-Ann MacDonald	University of Prince Edward Island
Dr. Mark Asbridge	Dalhousie University
Dr. Annie Montreuil	Université du Québec à Montréal
Dr. Tara Elton-Marshall	Centre for Addiction and Mental Health
Dr. Nathan Nickel	University of Manitoba
Dr. Nazeem Muhajarine	University of Saskatchewan
Dr. Cam Wild	University of Alberta
Dr. Marjorie MacDonald	University of Victoria

This manual has been produced to facilitate the manipulation and use of the 2018/2019 CSTADS Public Use Microdata File. Please refer to the 2018/2019 CSTADS Data Codebook and the 2018/2019 CSTADS Frequency File for related details.

1.1 Accessing Data and Recommended Acknowledgement

Access to the 2018/2019 CSTADS Public Use Microdata File (PUMF) is possible through Propel's Data Repository. Please take note of the data release guidelines found in Section 8.6 before using the data. The following text is a recommended acknowledgement in any publications / presentations arising from CSTADS data:

> *Data used for this research were taken from Health Canada's Canadian Student Tobacco, Alcohol and Drugs Survey (CSTADS), which was conducted for Health Canada by the Propel Centre for Population Health Impact at the University of Waterloo. Health Canada has not reviewed, approved, nor endorsed this research. Any views expressed or conclusions drawn herein do not necessarily represent those of Health Canada*.

PLEASE NOTE: **For information purposes only (and in confidence), Health Canada would appreciate receiving advance copies of planned publications arising from CSTADS data at least three (3) weeks prior to the publication date. Send a copy by mail or email to the Technical Authority noted above.**

2.0 Background

The Canadian Student Tobacco, Alcohol and Drugs Survey (CSTADS) serves as a key surveillance tool for Health Canada and its partners to help understand Canadian trends in student tobacco, alcohol and drug use. Understanding these trends is vital to the effective development and monitoring of policies and programs aimed at reducing substance use by this population (including the regulation of tobacco products and illicit drugs). It complements the general population Canadian Tobacco, Alcohol and Drugs Survey.

CSTADS collects data on tobacco use, alcohol and drug use among students in grades 7 through 12 (secondary I through V in Québec). CSTADS produces data comparable to that of the Youth Smoking Survey (YSS) (its former name), and provides Health Canada, its partners and stakeholders, as well as the Canadian public, with timely and reliable data on use of tobacco, alcohol and drugs by Canadian students. Additional data on bullying, and new data on sleep behaviours were gathered in the 2018/2019 cycle.

The biennial CSTADS is a school-based survey of a generalizable sample of schools in the ten Canadian provinces. When first administered in 1994, the YSS was the largest and most comprehensive survey on youth smoking behaviour since 1979. The YSS was administered to grades 5 through 9 until 2006/2007, when the coverage of the YSS was expanded to include Canadian youth in grades 5 to 12. As a result of low rates of tobacco use amongst grade 5 students, they were no longer included in the 2008/2009 and subsequent iterations of the YSS/CSTADS. The 2016/2017 cycle eliminated grade 6 students from the CSTADS sample.

All schools that participated in the 2018/2019 CSTADS, with the exception of those in New Brunswick (NB), received a $100 honorarium, a school-specific profile and two one-page summaries of their survey results within two months of the school's data collection. The summaries were targeted to school staff and administrators, students and the parent community, to facilitate sharing school-specific results. These school profiles and summaries provide valuable information for schools to address tobacco, alcohol and drug use, in addition to other social and school environment issues including bullying and sleep related behaviours. The profiles and summaries compared school-specific results to past provincial and national data, when available.

2.1 Collaborative Projects

The 2018/2019 CSTADS was implemented alongside two collaborative projects, one in Nova Scotia (NS) and one in New Brunswick (NB).

Nova Scotia Department of Health and Wellness

The Nova Scotia Department of Health and Wellness (NSDHW) requested that the usual CSTADS sample be expanded to obtain results generalizable at the health zone level in addition to the usual provincial level. This partnership resulted in a larger sample of CSTADS schools and students in NS. Sampling was adjusted to achieve generalizability within each of the four NS health zones. Twenty schools were added

to the NS school target number, making the CSTADS school target number in the province 44. See section 4.4 for more information on sampling in NS.

New Brunswick Student Wellness Survey

The New Brunswick Student Wellness Survey (NBSWS) is a school based census survey implemented by the New Brunswick Health Council (NBHC) every three years. The NBSWS collects data related to mental fitness, resilience, prosocial and oppositional behaviours, physical activity, healthy eating, tobacco-free living and bullying. To minimize school research burden the 2018/2019 CSTADS in NB was implemented in conjunction with the NBSWS. The majority of students completed the NBSWS, with every ninth student completing CSTADS. See section 4.4 for more information on sampling in NB.

3.0 Definitions Used in this Guide

This section describes terms and definitions used in this guide.

Approached: Sampled schools and their respective boards that were contacted to participate in the survey.

Boards Participating: A recruited school board with at least one participating school.

Eligible Class: All grade 7-12 classes within a participating school. Note: In rare cases, schools will limit the participating grades or number of participating classes.

Eligible Students: Students enrolled in an eligible class.

Participation Rate (%)[1]: The number of boards or schools that participated in the survey as a percentage of the total number of boards or schools, respectively, approached to participate in the project.

Refused: Boards or schools that refused to participate, did not respond, were unable to be reached, or withdrew their participation.

Response Rate (%): The number of grades 7 to 12 students who participated in the survey as a percentage of the total number of eligible students in participating grades 7 to 12 classrooms.

Sampled Schools: Schools selected from the sampling frame (described below). See section 4.0 for details on sampling methodology.

Sampling Frame: The list of eligible schools in Canada upon which the sample of schools for the project was drawn. Schools were eligible to be included in the sampling frame if the school had at least 20 students enrolled in any of the eligible grades (grades 7-12). Federally-funded schools, closed schools, schools for children with special needs, and First Nation reserve schools were not included in the sampling frame.

Schools Participating: Schools with at least one student questionnaire recorded in the final data file. The number of participating schools can exceed the number of targeted schools in a province.

---

[1] Note that in cycles of the project prior to 2012/2013, participation rate was referred to as recruitment rate in the Microdata User Guides

<u>Targeted # of Schools</u>: The number of participating schools the sampling design anticipates is required to achieve generalizable estimates to the provincial and national levels. The number of participating schools can exceed the number of targeted schools in a province. Please note that the targeted number of schools for Nova Scotia and New Brunswick reflect the collaborations with the Nova Scotia Department of Health and Wellness and the New Brunswick Health Council, respectively. See section 4.0 for details on sampling methodology.

4.0 Sampling Design

To obtain a sample of students, a stratified single stage cluster design was used. Schools were selected from strata, and then all eligible students within selected schools were surveyed. This process was used in all provinces except New Brunswick. See section 4.4 for details on sampling design in NB. The Propel School Database is updated annually and provides a current list of schools for the sampling frame.

4.1 Target Population

The target population for the 2018/2019 CSTADS consisted of all young Canadian residents attending private, public, and Catholic schools enrolled in one of grades 7 to 12 (secondary I to V in Québec), excluding schools in the Yukon, Northwest Territories, and Nunavut. Students who attend special schools (e.g., schools for visually-impaired, schools for hearing-impaired, daycares, special needs, First Nation reserve schools, virtual schools, schools located on military bases, international schools) were excluded from the target population. In addition, students from schools that do not have at least 20 students enrolled in at least one eligible grade were also excluded from the target population of schools to control data collection costs.

4.2 Sampling Stratification

In all participating provinces, except New Brunswick (NB), the sampling of the schools (and coincidentally, the boards within which each is situated) was based on a stratified single stage cluster design. The stratification was based on two classifications: health region[2] and type of school[3]. Within each provincial sampling frame, two to three health region categories and two school-level categories were defined. The allocation of the school sample to strata ensured adequate sample sizes inside the strata and increased precision.

Classification variable 1: Health Region Smoking Rate

In keeping with past cycles of YSS/CSTADS, stratifying by provincial health region smoking rate continued for 2018/19. Stratification by health region smoking rate ensures that schools located in health regions with high and low smoking rates are adequately represented. It also reduces the variability of smoking within strata, which increases the precision of the estimates for smoking status, amount smoked and other variables correlated with smoking (including other drug and alcohol use).

Within each province, schools were divided into two categories based on the smoking rate of 15-19 year olds in the health region in which the school was located. The school's six-digit postal code was used to identify the health region for a school.

---

[2] In Prince Edward Island there is only one Health Region, therefore there is only one health region category for Prince Edward Island.

[3] In Québec secondary schools comprise of grades secondary I to secondary V. Since these are the only grades we will collect data from, Québec will only have one school-level category: secondary.

The smoking rate of 15-19 year olds in each health region was based on the 2013/2014 Canadian Community Health Survey data, which were the most current data available at the time of sampling.

The goal was to create two groups of schools per province, each group homogenous in terms of health region smoking rates, with similar numbers of schools in each. To accomplish this, the smoking rates of 15-19 year olds from the Canadian Community Health Survey for each health region in a province were calculated and compared to each other within that province. Schools located in a health region with a smoking rate lower than the median smoking rate for the province were assigned to the "low" smoking rate health-region category. Similarly, schools located in a health region with a smoking rate greater than the median was assigned to the "high" smoking rate health-region category. Schools located in the health region with the median smoking rate were assigned to the "high" smoking rate health-region category provided this resulted in "low" and "high" categories with fairly similar numbers of schools. Occasionally, when there was a large number of schools in the health region with the median rate, the percentage of schools assigned to the "high" category compared to the "low" category became very high. When this occurred and presented the danger of having too few schools available to meet targets, schools in the health region with the median rate were re-assigned to the "low" category.

As a result of implementing previous cycles of YSS/CSTADS and to ensure adequate representation of schools in the most highly populated areas of certain provinces, a third "urban" category was defined in Newfoundland and Labrador (NL), Québec (QC), Ontario (ON) and Alberta (AB). Two categories paralleled the health region categories (high and low) described above, but excluded schools defined as being part of the urban areas of St. John's in NL, Montréal in QC, Toronto in ON, and Calgary/Edmonton in AB.

The urban category acknowledges the size of metropolitan areas in QC, ON and AB and was developed as a result of the difficulties experienced in past cycles of CSTADS (YSS) related to school board refusals and lower student permission rates in these specific urban areas.

Sampling schools located in large urban centres ensured representation from these centres. In NL, the third category consisted of all schools in the Eastern Regional Health Authority. The urban category was implemented in NL because the Eastern Health Region, which is the largest of the health regions in this province, had the median smoking rate. Assigning the Eastern Health Region to either the "low" or the "high" smoking rate health-region category resulted in very skewed "low"/"high" categories making school targets for the low category impossible. The only way to ensure even health region categories was to include an urban category in NL. In QC, the urban stratum is defined as all schools listed within the Greater Montréal area which includes the entire Island of Montréal, Laval and the Urban Agglomeration of Longueuil. In ON, the third category consisted of all schools in the Greater Toronto Area (GTA), defined as comprising these health units: Toronto Regional Health Unit, York Regional Health Unit, Peel Regional Health Unit, Halton Regional Health Unit and Durham Regional Health Unit. In AB, the third category consisted of all schools located within 20 kilometers of Calgary and 20 kilometers of Edmonton, included schools located in Calgary, Edmonton, St. Albert and Sherwood Park.

Classification variable 2: School Type

Schools were defined as being a member of either an elementary[4] or secondary[5] school category. When the total enrolment of elementary grades was greater than or equal to the total enrolment of the secondary grades for a school, the school was assigned to the "elementary" school category. Similarly, when the total enrolment of secondary grades was greater than the total enrolment of elementary grades, the school was assigned to the "secondary" school category.

4.3   Sample Selection

Crossing these classification variables yielded six strata in Newfoundland and Labrador, Ontario and Alberta, four strata in British Columbia, Saskatchewan, and Manitoba, three in Québec, and two in Prince Edward Island. Within each stratum, in each province, schools were selected based on simple random sampling. Typically, two elementary schools for every one secondary school were sampled to ensure appropriate distribution of participants across all grades, given that the elementary schools characteristically have lower enrolments than secondary schools.

The number of schools sampled in each province ensures a minimum sample that would produce generalizable results, and is based on the elementary and secondary school participation rates in previous cycles of YSS/CSTADS. The participation rate is defined as the total number of schools that participated in the survey divided by the number of schools sampled. Once a sampled list of schools for a province is generated, the co-investigator leading the CSTADS implementation in that province reviews the list and comments on the overall distribution of schools and identifies any problems with the sample that may be of concern at the provincial level.

All students within eligible grades in each sampled school are invited to participate in the survey. Schools have indicated that, administratively, surveying their entire school population is simpler both in terms of obtaining parental permission and questionnaire implementation.

Simple random sampling of schools within a stratum, and then surveying all eligible grades in each selected school, implies equal inclusion probabilities at the outset for all students in the stratum.

4.4   CSTADS Sample Selection when Partnering with Provincial Surveys

New Brunswick Sampling Design

The New Brunswick Student Wellness Survey (NBSWS) is a near census survey. All New Brunswick schools with at least one grade within 6 to 12 were invited to participate in the collaborative NBSWS-CSTADS study. Schools with a total enrollment of 50 students or less were excluded from the CSTADS sample of schools. Exclusion of these smaller schools ensured that school level feedback results provided by the New Brunswick Health Council are not compromised due to low student participant numbers. Within each eligible school, questionnaires were distributed among students using an 8:1 ratio of distribution with the majority of students receiving the NBSWS and the remaining the CSTADS questionnaire. Students in grade 6 participated in the NBSWS but were excluded from CSTADS.

---

[4] Elementary grades defined as grades 7 to 9 in Alberta and grades 7 to 8 in all other provinces.

[5] Secondary grades defined as grades 10 to 12 in Alberta, Secondary I to V (grades 7 to 11) in Québec, and grades 9 to 12 in all other provinces.

CSTADS in Nova Scotia sampling design

For the 2018/2019 cycle, Propel partnered with the Nova Scotia Department of Health and Wellness (NSDHW). Nova Scotia was interested in obtaining data from each NS health zone. To accommodate, CSTADS in NS was implemented by stratifying by health zone and adding extra schools to the sample to obtain generalizable data within each zone. Therefore, in NS schools were stratified by school type (elementary, secondary) and health zone (four zones), yielding eight strata.

5.0 Questionnaire Development

The final 2018/2019 CSTADS questionnaire (see Appendix A) was developed through a series of meetings and pilot testing sessions. Appendix B lists questions that have appeared in all cycles of the CSTADS (YSS) since 1994 and a separate list of questions that are new to the 2018/2019 CSTADS.

Several key considerations guided the design and content of the questionnaire:

- **Comparability**: The basis of the questionnaire was past versions of the CSTADS (YSS) questionnaire (2002-2016), to allow for comparisons across cycles.

- **Responsiveness**: To meet the needs of users of the data, CSTADS investigators and those responsible for federal and provincial tobacco, alcohol and drug use strategies were given an opportunity to contribute topics/items for consideration.

- **Relevancy**: To ensure value-added for participating schools, items and content areas (i.e., bullying, mental well-being, and school connectedness) were added to enhance the school-specific results profiles and summaries to schools.

- **Feasibility**: To meet the criterion of students being able to complete the questionnaire in one class period (30 minutes), questionnaire length was restricted.

The final content of the questionnaire included questions on demographics, tobacco use, alcohol and drug use, bullying, and sleep related behaviours. The questionnaire was available in English and French (see Appendix A). Copies of past CSTADS (YSS) questionnaires can be found at www.cstads.ca

5.1 Pilot Testing

Two rounds of questionnaire pilot testing occurred prior to implementing the 2018/2019 CSTADS in schools: English-language (in Stratford, Ontario), and French-language (in Montréal, Québec)[6].

In each setting, student groups (4 English, 2 French) completed the CSTADS questionnaire and participated in focus groups immediately after. The pilot testing established student understanding of items, and gathered feedback from the students on new questions, questionnaire flow and overall completion times. A number of modifications were made to the questionnaire as a result of the pilot testing. The French language used in the questionnaire was deemed suitable by all French- language participants.

_____

[6] University of Waterloo. 2018 (April).Canadian Student, Tobacco, Alcohol and Drugs Survey (CSTADS): 2018/2019 CSTADS English and French Pilot Tests Report. Waterloo: Propel Centre for Population Health Impact, University of Waterloo. 1-41. Available upon request.

6.0 Survey Implementation

6.1 Ethics Review

The 2018/2019 CSTADS project required ethics approval from the University of Waterloo Office of Research Ethics, the Health Canada Research Ethics Board, the ethics review boards located in affiliated provincial institutions, and school board ethics review committees. The CSTADS team obtained ethics approval from all research ethics boards noted above prior to implementation.

All amendments, modifications and adverse events were reported to the appropriate review committees (as needed). No major incidents occurred during the implementation of the 2018/2019 CSTADS in schools across the country. All minor incidents were satisfactorily resolved.

6.2 Data Quality: Recruitment and Response Rates

CSTADS implementation includes three levels of recruitment. Within each province, the team recruits school boards (divisions/districts), schools, and then students, to participate. As a result, non-response occurs at various levels. Refer to Tables 1, 2 and 5 for response rates by province at the school board, school and students by grade, respectively.

Board and School Recruitment and Participation

Trained teams within each province took responsibility for recruiting all boards and schools within their province. Schools with no governing school board were approached directly. After successfully recruiting school boards, randomly selected schools were approached. An online, real-time database used by all staff permitted central monitoring by Propel to ensure that protocols and progress were consistent across the country.

A total of 116 school boards and 442 schools participated in the 2018/2019 CSTADS. See Tables 1-4 for details on board and school participation by province and sampling stratum.

Table 1: Board Participation Outcomes by Province, 2018/2019 CSTADS

Province	# Boards Approached	# Boards Recruited	# Boards Refused	# Boards Participated^
NL	1	1	0	1
PE	2	2	0	2
NS	8	6	2	6
NB	7	7	0	7
QC	38	30	8	13
ON	55	38	17	31
MB	27	12	15	11
SK	27	11	16	9
AB	48	28	20	18
BC	44	30	14	18
Canada	**257**	**165**	**92**	**116**

^Recruited school boards with at least one participating school.

Table 2: School Participation Outcomes by Province, 2018/2019 CSTADS

Province	# Target Schools	# Sampled Schools	# Schools Approached	# Schools Recruited	# Schools Participated	# Schools Refused
NL	24	36	29	24	24	5
PE	24	30	25	22	21	3
NS	44	109	74	50	44	24
NB	151	151	151	148	139	3
QC	20	109	66	28	25	38
ON	54	369	153	66	60	87
MB	28	93	44	29	28	15
SK	36	375	78	40	37	38
AB	32	156	62	34	32	28
BC	32	172	91	37	32	54
Canada	**445**	**1600**	**773**	**478**	**442**	**295**

Table 3: Participating Schools by School Categories and Province, 2018/2019 CSTADS

Province	School Category	Target # Schools	# Schools Participated
NL	Elementary	12	12
	Secondary	12	12
PE	Elementary	16	13
	Secondary	8	8
NS	Elementary	24	24
	Secondary	20	20
QC	Secondary	20	25
ON	Elementary	36	39
	Secondary	18	21
MB	Elementary	18	17
	Secondary	10	11
SK	Elementary	18	18
	Secondary	18	19
AB	Elementary	20	20
	Secondary	12	12
BC	Elementary	20	20
	Secondary	12	12
Canada		**294**	**303**

Table 4: Participating Schools by Health Region Smoking Rate Categories and Province, 2018/2019 CSTADS

Province	Total Target # Schools	Smoking Rate Category	Target # Schools by Smoking Rate	# Schools Participating
NL	24	Low	6	6
		High	8	8
		Urban	10	10
PE	24	n/a	24	21
NS	44	Region 1	11	11
		Region 2	11	11
		Region 3	11	12
		Region 4	11	10
QC	20	Low	6	8
		High	6	8
		Urban	8	9
ON	54	Low	18	22
		High	18	19
		Urban	18	19
MB	28	Low	14	14
		High	14	14
SK	36	Low	18	21
		High	18	16
AB	32	Low	10	10
		High	10	10
		Urban	12	12
BC	32	Low	16	16
		High	16	16
Canada	**294**		**294**	**303**

Student Recruitment and Participation

Within recruited schools, all students in grades 7 to 12 were invited to participate in the survey. Only students with parental permission (or of age to consent to participate) participated in the survey. A total of 62,850 students across 10 Canadian provinces participated. See Table 5 for student participation details by grade and province. A generalizable sample was achieved at the national level and at the provincial level for all participating provinces.

A mix of active permission protocols (active) and active information-passive permission (passive) methods were used to obtain parent permission in participating schools. School boards and schools determined the permission method most appropriate within their schools.

Schools sent information and permission materials home to parents of students enrolled in grade 7 to 12 classes. The information letters provided details about the project, contact information for project staff, and referred parents to the project website for further details and copies of the questionnaires. Parents had a minimum of two weeks to respond. Schools participating with active protocols required that only those students whose parents indicated "yes" on a permission form were able to participate in

the survey. Schools participating with passive permission protocols required that parents call a toll-free number if they did not want their child to participate in the survey. Students also had the opportunity to decline participation on the day of data collection.

Table 5: Participating Students by Province and Grade, 2018/2019 CSTADS

Province	Number of Students Participating, by Grade						
	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Total
NL	1067	901	860	922	779	792	5321
PE	993	936	1063	628	615	519	4754
NS	1309	1285	1097	1131	889	759	6470
NB	534	585	599	580	612	567	3477
QC	2924	2782	3516	3458	3163	0	15843
ON	1690	1567	930	866	494	516	6063
MB	749	770	594	532	467	406	3518
SK	877	804	635	604	463	485	3868
AB	1065	943	1167	1335	956	852	6318
BC	1362	1478	1293	1206	1057	822	7218
Canada	12570	12051	11754	11262	9495	5718	62850

Approximately 84% of students participated with passive permission and 16% of students participated with active permission, excluding NB. Among participating classes, non-response at the student level can be attributed to several factors, including parent/guardian refusals and student refusals or absenteeism on the day of survey implementation. The response rate at the student level is derived based on the number of eligible students as provided by school contacts for participating classes. Every effort is made to recruit all eligible classes in participating schools but on occasion, some eligible classes do not participate. See Table 6 for student level response rates.

Table 6: Student Level Response Rates by Province, 2018/2019 CSTADS

Province	Eligible students	Students with active permission	Students with passive permission	Completed questionnaires	Response rate [+]
NL	6579	0	5321	5321	81%
PE	5586	52	4702	4754	85%
NS	12824	2976	3494	6470	50%
NB++	4812	-	-	3477	72%
QC	17584	0	15843	15843	90%
ON	14570	3879	2184	6063	42%
MB	4668	257	3261	3518	75%
SK	4775	23	3845	3868	81%
AB	10202	2019	4299	6318	62%
BC	9071	376	6842	7218	80%
Canada	90671	9582	49791	62850	69%

[+] Based on completed questionnaires (numerator) and eligible students (denominator) in participating classes.

++The number of eligible students in NB is 1/9 of all eligible students in schools participating in both the NBSWS and CSTADS. Numbers of students with active/passive permission in NB are not available due to the collaboration in the province.

6.3   Classroom Implementation

Data collections occurred between October 29, 2018 and June 17, 2019. On the day of a school's data collection, classroom teachers followed standard and simple project instructions to administer the paper questionnaire during a designated class period. The questionnaire administration, including instructions to the students, required 40 minutes or less within each class. To protect confidentiality, teachers were asked not to circulate within the classroom while students were completing the questionnaire, and each student put their completed questionnaire in a sealable envelope before it was collected by a fellow student. In most cases, CSTADS staff were available at the school on the day of data collection for support. Exceptions included remote schools.

7.0 Creating the Public Use Microdata File (PUMF)

CSTADS produces a Public Use Microdata File (PUMF) in SAS, SPSS and STATA that can be requested from the Propel's Data Repository (https://uwaterloo.ca/propel/resources-and- products/propels-data-repository). This section presents a brief summary of the steps involved in producing the PUMF and considerations required when conducting analyses using the PUMF.

Note: The system for naming the variables changed in the 2014/2015 CSTADS PUMF. This resulted in new variable names for variables that existed on previous CSTADS (YSS) PUMF files. Please take this into account when analysing data over time. See Appendix B-1 for a listing of select past (since 2004) and current variables names used in the CSTADS (YSS) data files.

7.1   Data Capture

Completed questionnaires were machine-scanned using Optical Mark Recognition (OMR) technology. Procedures detailed several quality-control measures to ensure the accuracy of the scanned data. Processing staff visually scanned all questionnaires to ensure that the OMR would correctly record the data on the questionnaires.

In the course of visually scanning a questionnaire, processing staff could "correct" a questionnaire according to defined rules, including: darkening marks that were too light or incomplete (e.g., check marks instead of filled-in circles); erasing marks from answers where the respondents changed their mind but did not sufficiently erase the original response; erasing accidental/wayward marks that were not meant to indicate answers (e.g., graffiti or doodles); and, erasing marks made in any places reserved for "office use only". If processing staff were unsure of how to proceed with an answer, the project manager would instruct them and/or would seek the advice of a data analyst as necessary.

7.2   Editing and Imputation

The Public Use Microdata File (PUMF) uses the following standard codes:

Valid skip - 96 and 996

Not stated - 99 and 999

Prior to data cleaning, the 2018/2019 CSTADS student dataset contained 63,250 records and retained 62,850. Record removal occurred when respondents did not record sex (382 records), did not record grade and the grade could not be imputed (16 records, 11 of which also did not record sex) or completed only demographic questions (13 records). In addition, Québec secondary I, II, III, IV, and V were converted to grades 7, 8, 9, 10, and 11, respectively.

Note to SPSS users: in the SPSS PUMF, many variables have the values 96 and/or 99, which are defined as "Missing". These cases are automatically excluded from the analysis when producing estimates for these variables. These cases can be changed by the user in the "Missing" column in the SPSS "Variable View" of the dataset.

Treatment of the 'I do not know' or 'I do not remember' Option

There are several variables in the dataset that require attention to the "I do not know" option. We recommend treating the "I do not know" response as a "Not Stated (99 or 999)" response for the following questions:

- SS_020 (Question 8) ("How old were you when you first tried smoking cigarettes, even just a few puffs?")

- PP_021 (Question 19) ("Thinking about the last time you bought cigarettes in the last 12 months, how much did you pay?")

- ELC_024 (Question 22) ("On how many of the last 30 days did you use an e-cigarette (vape, vape pen, tank & mod)?")

- ALC_020 (Question 26) ("In the last 12 months, how often did you have a drink of alcohol that was more than just a sip?")

- ALC_030 (Question 27) ("How old were you when you first had a drink of alcohol that was more than just a sip?")

- ALC_040 (Question 28) ("In the last 30 days, how often did you have a drink of alcohol that was more than just a sip?")

- ALC_050 (Question 29) ("In the last 12 months, how often did you have 5 or more drinks of alcohol on one occasion?")

- ALC_060 (Question 30) ("How old were you when you first had 5 drinks or more of alcohol on one occasion?")

- ALC_071 (Question 31) ("In the last 30 days, how often did you have 5 or more drinks of alcohol on one occasion?")

- CAN_020 (Question 34) ("In the last 12 months, how often did you use marijuana or cannabis?")

- CAN_030 (Question 35) ("How old were you when you first used marijuana or cannabis?")

- CAN_040 (Question 36) ("In the last 30 days, how often did you use marijuana or cannabis?")

- CAN_140 (Question 40) ("In the last 12 months, how often did you have alcohol AND marijuana or cannabis on the same occasion? (e.g., at a party, in the same evening, etc.)")

<u>Note to SPSS Users</u>: In the SPSS version of this file, 'I do not know' and 'I do not remember' are defined as "Missing" (along with values 96 and 99) and will not be included in estimates produced using these variables. These cases can be changed by the user in the "Missing" column in the SPSS Variable View of the dataset.

Treatment of the 'Other' Option

No cleaning was done to the response options 'Other' in all variables with this type of response option, with the exception of gender and sexual orientation (questions 4 and 5), which are not in the PUMF in order to maintain confidentiality. Written answers were not recorded. The following is a comprehensive list of all variables that include the 'Other' option:

- CA_011 (Question 17) ("Where do you usually get your cigarettes?")
- ELC_050j (Question 23) ("Where do you usually get your e-cigarettes and supplies (vape, vape pen, tank & mod, e-juice)?")
- CAN_121 (Question 38) ("In the last 12 months how did you usually get the marijuana or cannabis you used?")
- PR_110 (Question 51) ("In the last 12 months if you did use prescribed pain relievers for non-medical reasons or to get high, how did you get them?")

Grade (Question 1)

During the cleaning process, data analysts investigated cases where grade was missing or inconsistent with the grades represented in their school. If a student indicated a grade that did not match the relevant grades in the school or province, or if grade was missing or uncodeable, the variable was recoded to the median grade of the class to which the student belonged. If this option was not available, then student age was used to impute grade.

Age (Question 2)

If age was out of range for grade, then age was set to "99 = Not Stated". To make every effort to maintain confidentiality, self-reported age (question 2) is not included in the PUMF.

Gender and Sexual Orientation (Questions 4 and 5)

The gender (question 4) and sexual orientation (question 5) questions each gave the respondent an opportunity to write in an open ended response if they felt that they did not fit into any of the listed categories. The open ended responses of all respondents who chose either "Or please specify" or "Something else. I identify as" were categorized as either recode to one of the options listed, legitimate other, pansexual (a very prevalent "legitimate other" in question 5 only), non valid response, "I do not understand the question" (question 4 only), or "Wrote Normal as their sexual orientation" (question 5 only). These categorizations were then applied to questions 4 and 5, in order to create the derived variables DVGENDER and DVORIENT where those who chose "Or please specify" or "Something else. I identify as" were recoded as one of the options above or they were given a "99 = Not Stated" if they did not include a written response. To make every effort to maintain confidentiality, all variables related to self-reported gender and sexual orientation (questions 4 and 5) are not included in the PUMF.

Age of Initiation Variables

If a reported age of initiation was greater than the student's self-reported age (question 2), then age of initiation was assigned a value of "99 = Not Stated". If the self-reported age (question 2) was "99 = Not

Stated", then the reported age of initiation was compared to an imputed version of age based on median age per class and the grade. If the imputed age was found to be less than the age of initiation reported then the age of initiation was given a value of "99 = Not Stated". This applies to the following age of initiation questions:

- SS_020 (Question 8) ("How old were you when you first tried smoking cigarettes, even just a few puffs?")
- ALC_030 (Question 27) ("How old were you when you first had a drink of alcohol that was more than just a sip?")
- ALC_060 (Question 30) ("How old were you when you first had 5 drinks or more of alcohol on one occasion?")
- CAN_030 (Question 35) ("How old were you when you first used marijuana or cannabis?")

NEW: Question 15 (WP_040a to WP_040g) ("Thinking back over the last 7 days, how many whole cigarettes did you smoke each day?")

This one question converts to 7 variables, one for each day of the week. The range that was allowed for each variable was 0 - 36. All responses of 37 or greater have been set to "99 = Not Stated". Respondents who have not smoked a whole cigarette (SS_030 = 2 or 96) were set to "96 = Valid Skip" for each of these variables.

Imputations of Core Smoking Questions

As part of the CSTADS data cleaning we impute missing values for the core smoking questions SS_010 (Question 7), SS_030 (Question 11), SS_040 (Question 12), and SS_060 (Question 13). Historically this was done by applying either an algorithm or the SAS procedure for imputing missing variables (PROC MI) and supplementing this approach by checking imputed records with seemingly inconsistent responses. For CSTADS 2018/2019, we used the algorithm developed for CSTADS 2016/2017. This strategy was meant to ensure that any imputations were based on the collective evidence of each respondent's smoking mode and not on the probability of a respondent's habits based on other respondents' patterns. When records did not satisfy the algorithm, inconsistent and missing responses were left as provided by the student.

1) An algorithm based on the core questions described above as well as other questions in the survey, SS_020 (Question 8), WP_030 (Question 14), WP_040a to WP_040g (Question 15), SC_010 (Question 16), and CI_010 (Question 21) was applied to impute missing values of the core smoking questions. The development of the algorithm occurred last cycle (2016/2017) and is different from the original algorithm as well as the other methods used in previous cycles. One of the main differences is that some records remain missing where it did not make sense to impute. In cycles prior to 2016/2017 all missing records were imputed.

2) We completed the imputations before the application of skip patterns (described below) and the skip patterns can overwrite imputed values. For example, if SS_010 (Question 7: Have you ever tried cigarette smoking, even just a few puffs?) is "No" and SS_030 (Question 11: Have you ever smoked a whole cigarette?) is missing/"not stated" before it was sent through the imputation procedure, the "No" from SS_010 would dictate that SS_030 will be 96 (valid skip) in the final data regardless of the outcome of the imputation.

Question 23 (ELC_050a to ELC_050j)

If the student had selected "I do not use e-cigarettes" plus at least one of the other response options, all of these variables were equated to "99=Not Stated".

With the exception of the situation described above, if at least one item was selected then the variable for that item was coded to "1=Yes" and all other variables that were not selected were coded to "2=No". These variables were coded to "99=Not Stated" only when the whole question had been left blank or when there were inconsistencies as described in the paragraph above.

Skip Patterns and Inconsistent Responses

The questionnaire was intentionally designed with no respondent-use skip patterns and included response options like "I do not smoke" and "I did not use". These response options help avoid the identification of substance users by rate of questionnaire completion time in the classroom. Skip patterns are imposed on many follow-up variables after an initial question has established whether the respondent is a user or non-user of a given substance (tobacco, alcohol, drugs).

Unlike a computer-assisted questionnaire, which may be programmed to generate prompts when inconsistent responses are provided (e.g., age at first cigarette is older than current age), there can be no built-in verification of answers in a paper-and-pencil survey. Inconsistent answers provided by respondents are retained in the data file except where skip patterns or algorithms have been imposed during the preparation of the PUMF. Imposed skip patterns overwrite all related dependent answers provided by respondents.

Starting with the 2016/2017 cycle and continuing in the 2018/2019 cycle, the core smoking questions have some missing values left as is. Skip patterns were not applied based on non-response for the tobacco questions. To have a valid skip applied, the response to the previous question must be either "No" or "Valid Skip".

Data cleaning procedures in 2018/2019 and 2016/2017 were unchanged from previous cycles for all drug and alcohol questions.

Please refer to the 2018/2019 CSTADS Data Codebook for details on the coverage for questions where skip patterns were imposed.

7.3   Derived Variables

The PUMF includes derived variables, created by combining questionnaire items, to facilitate data analysis and ensure consistency across users. The following describes the derived variables included in the 2018/2019 CSTADS PUMF. Four derived variables previously included in the 2014/2015 CSTADS (DVAMTSMK, DVCIGWK, DVNDSMK, DVAVCIGD) and which were not included in the 2016/2017 CSTADS PUMF, were computed using the number of cigarettes smoked each of the last 7 days (WP_040a to WP_040g). These derived variables were omitted in 2016/2017 and are now being reintroduced because the source question(s) were different for 2016/2017 only.

Seven derived variables previously included in the 2016/2017 CSTADS will not be included in the 2018/2019 CSTADS since their component questions are not included on the 2018/2019 questionnaire: DVSUSCEP, DVFLAV30, DVRELATE, DVAUTON, DVCOMP, DVSOC_P and DVSOC_N.

Derived Variables for Smoking Status

The following describes the derived variables in the PUMF for smoking status. These derived variables are consistent with previous cycles of CSTADS (YSS).

DERIVED VARIABLE	DVTY1ST
RESPONSE OPTIONS FOR DVTY1ST	1 = Current Smoker
	2 = Former Smoker
	3 = Never Smoker
	99 = Not Stated
COVERAGE FOR DVTY1ST	Coverage: All Respondents
QUESTIONS	SS_030 (Question 11): *Have you ever smoked a whole cigarette?*
	1 (Yes)
	2 (No)
	96 (Valid Skip)
	99 (Not Stated)
	Coverage: All Respondents where SS_010 = 1 or 99 (Ever tried smoking a cigarette, even a few puffs or not stated)
	SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?*
	1 (Yes)
	2 (No)
	96 (Valid Skip)
	99 (Not Stated)
	Coverage: All Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)
	SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?*
	1 (None)
	2 (1 day)
	3 (2 to 3 days)
	4 (4 to 5 days)
	5 (6 to 10 days)
	6 (11 to 20 days)
	7 (21 to 29 days)
	8 (30 days (every day))
	96 (Valid Skip)
	99 (Not Stated)
	Coverage: All Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)

DERIVATION OF RESPONSES FOR DVTY1ST

CURRENT SMOKER	**Definition**:
	A current smoker is someone who has smoked at least 100 cigarettes in his or her lifetime, and who has smoked at least one whole cigarette during the past 30 days
	Calculation:
	SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 1 (Yes)
	AND
	SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* Valid responses 2 (1 day) 3 (2 to 3 days) 4 (4 to 5 days) 5 (6 to 10 days) 6 (11 to 20 days) 7 (21 to 29 days) 8 (30 days (every day))
FORMER SMOKER	**Definition:**
	A former smoker is a person who reports having smoked 100 or more cigarettes but did not smoke in the last 30 days.
	Calculation:
	SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 1 (Yes)
	OR
	SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* Valid response 1 (None)

NEVER SMOKER	**Definition:**
	A never smoker is a person who reports that he or she has not smoked 100 or more whole cigarettes in his or her life time but might have smoked a whole cigarette.
	Calculation:
	SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 2 (No) OR SS_030 (Question 11): *Have you ever smoked a whole cigarette* Valid responses 2 (No) 96 (Valid Skip)
NOT STATED	**Definition:**
	Smoking status unknown. Respondents whose smoking status could not be determined due to missing responses in one or more contributing variable.
	Calculation:
	Remaining respondents who were not assigned a smoking status based on the logic above.

DERIVED VARIABLE	DVTY2ST
RESPONSE OPTIONS FOR DVTY2ST	1 = Current Daily Smoker
	2 = Current Occasional Smoker
	3 = Former Smoker
	4 = Experimental Smoker (Beginner)
	5 = Past Experimental Smoker
	6 = Puffer
	7 = Never Tried
	99 = Not Stated
COVERAGE FOR DVTY2ST	Coverage: All Respondents
QUESTIONS	SS_010 (Question 7): *Have you ever tried cigarette smoking, even just a few puffs?*
	1 (Yes)
	2 (No)
	99 (Not Stated)
	Coverage: All Respondents
	SS_030 (Question 11): *Have you ever smoked a whole cigarette?*
	1 (Yes)
	2 (No)
	96 (Valid Skip)
	99 (Not Stated)
	Coverage: All Respondents where SS_010 = 1 or 99 (Ever tried smoking a cigarette, even a few puffs or not stated)
	SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?*
	1 (Yes)
	2 (No)
	96 (Valid Skip)
	99 (Not Stated)
	Coverage: All Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated.
	SS_060 (Question 13): On how many of the last 30 days did you smoke one or more cigarettes?
	1 (None)
	2 (1 day)
	3 (2 to 3 days)
	4 (4 to 5 days)
	5 (6 to 10 days)

6 (11 to 20 days)
7 (21 to 29 days)
8 (30 days (every day))
96 (Valid Skip)
99 (Not Stated)

Coverage: All Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)

DERIVATION OF RESPONSES FOR DVTY2ST	
CURRENT DAILY SMOKER	**Definition:** A current daily smoker is a person who reports currently smoking cigarettes every day. **Calculation:** SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 1 (Yes) AND SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* Valid response 8 (30 days (every day))
CURRENT OCCASIONAL SMOKER	**Definition:** A current occasional smoker is a person who currently smokes cigarettes but not every day. **Calculation:** SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 1 (Yes) AND SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* Valid responses 2 (1 day) 3 (2 to 3 days)

	4 (4 to 5 days) 5 (6 to 10 days) 6 (11 to 20 days) 7 (21 to 29 days)
FORMER SMOKER	**Definition:** A former smoker is a person who smoked at least 100 cigarettes in his/her lifetime but did not smoke at all in the last 30 days. **Calculation:** SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 1 (Yes) AND SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* Valid Response 1 (None)
EXPERIMENTAL SMOKER (BEGINNER)	**Definition:** An experimental smoker is a person who has smoked in the last 30 days but has not smoked 100 or more cigarettes. **Calculation:** SS_030 (Question 11): *Have you ever smoked a whole cigarette?* Valid response 1 (Yes) AND SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 2 (No) AND

	SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarette*s?
	Valid responses 2 (1 day) 3 (2 to 3 days) 4 (4 to 5 days) 5 (6 to 10 days) 6 (11 to 20 days) 7 (21 to 29 days) 8 (30 days (every day))
PAST EXPERIMENTAL SMOKER	**Definition:** A past experimental smoker is a person who has smoked a whole cigarette but did not smoke in the last 30 days and also did not smoke 100 cigarettes in his/her lifetime. **Calculation:** SS_030 (Question 11): *Have you ever smoked a <u>whole</u> cigarette?* Valid response 1 (Yes) AND SS_040 (Question 12): *Have you ever smoked 100 or more whole cigarettes in your life?* Valid response 2 (No) AND SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* Valid response 1 (none)
PUFFER	**Definition:** A puffer is a person who has tried smoking, but has never smoked a whole cigarette. **Calculation:** SS_010 (Question 7): *Have you <u>ever</u> tried*

	cigarette smoking, even just a few puffs? Valid response 1 (Yes) AND SS_030 (Question 11): *Have you ever smoked a <u>whole</u> cigarette?* Valid response 2 (No)
NEVER TRIED	**Definition:** A person classified as never tried, has never tried a cigarette, not even just a few puffs. **Calculation:** SS_010 (Question 7): *Have you <u>ever</u> tried cigarette smoking, even just a few puffs?* Valid response 2 (No)
NOT STATED	**Definition:** Detailed smoking status unknown. Respondents whose detailed smoking status could not be determined due to missing responses in one or more contributing variable. **Calculation:** Remaining respondents who were not assigned a smoking status based on the logic above.

Derived Variable for Smoking in the Last 30 Days

The following describes the derived variable in the PUMF for smoking in the last 30 days.

DERIVED VARIABLE	DVLAST30
OBJECTIVE	To measure the rate of having smoked one or more cigarettes in the last 30 days out of all respondents.
RESPONSE OPTIONS FOR DVLAST30	1= Yes 2= No 99 = Not Stated
COVERAGE FOR DVLAST30	Coverage: All Respondents
QUESTIONS	SS_060 (Question 13): *On how many of the last 30 days did you smoke one or more cigarettes?* 1 = None 2 = 1 day 3 = 2 to 3 days 4 = 4 to 5 days 5 = 6 to 10 days 6 = 11 to 20 days 7 = 21 to 29 days 8 = 30 days (every day) 96 = Valid Skip 99 = Not Stated Coverage: All Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated) Notes: • Code "Yes" for respondents who reported 1-30 days of cigarette use in question 13 (2 <= SS_060 <= 8). • Code "No" for respondents who did not report 1-30 days of cigarette use in question 13 (SS_060 = 1 (None) or 96 (Valid Skip)). • Code "Not Stated" for respondents who did not respond to question 13 (SS_060 = 99).

Derived Variables for Amount Smoked

The following describes the derived variables in the PUMF for amount smoked. These derived variables are consistent with previous cycles of CSTADS (YSS) of 2014/2015 and earlier.

COMPONENT VARIABLES	WP_040A, WP_040B, WP_040C, WP_040D, WP_040E, WP_040F, WP_040G
QUESTION	(All seven days from Question 15) *Thinking back over the last 7 days, how many whole cigarettes did you smoke each day?* *a) Sunday* *b) Monday* *c) Tuesday* *d) Wednesday* *e) Thursday* *f) Friday* *g) Saturday*
RESPONSE OPTIONS	0 = 0 whole cigarettes smoked 1 : 36 whole cigarettes smoked 96 = Valid Skip 99 = Not Stated
COVERAGE	Coverage: Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)

DERIVED VARIABLES

DVAMTSMK	**Definition**: The average number of whole cigarettes smoked per day across all 7 days in the past week. **Calculation of Responses:** (WP_040a + WP_040b + WP_040c + WP_040d + WP_040e + WP_040f + WP_040g)/7 **Notes:** If any component variables have missing data then DVAMTSMK = 99. **Coverage:** Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)
DVCIGWK	**Definition**:

Total number of whole cigarettes smoked in the last 7 days.

Calculation of Responses:
WP_040a + WP_040b + WP_040c + WP_040d + WP_040e + WP_040f + WP_040g

Notes:
- Not necessary for all to have valid responses.
- If all component variables have missing data then DVCIGWK = 999.

Coverage:
Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)

DVNDSMK	**Definition**: Number of days the respondent smoked at least one whole cigarette in the week prior to the survey. **Calculation of Responses:** A count of WP_040a, WP_040b, WP_040c, WP_040d, WP_040e, WP_040f, and WP_040g with valid responses excluding days with a missing or zero response. **Notes:** - If all component variables are equal to the numeric value 0 or if some of the component variables are equal to the numeric value 0 and some are missing then DVNDSMK = 0. - If all component variables have missing data then DVNDSMK = 99. **Coverage:** Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)
DVAVCIGD	**Definition**: Average number of whole cigarettes smoked on the days that the respondent smoked.

Calculation of Responses:
DVCIGWK / DVNDSMK

Notes:
- If DVCIGWK and DVNDSMK = 0 then DVAVCIGD = 0.
- If either DVCIGWK or DVNDSMK were missing, then DVAVCIGD = 99.

Coverage:
Respondents where SS_030 = 1 or 99 (Ever smoked a whole cigarette or not stated)

Derived Variable for Number of years lived in Canada

The following describes the derived variable in the PUMF for the number of years lived in Canada.

DERIVED VARIABLE	DVRES
OBJECTIVE	To measure the number of years the respondent lived in Canada. Categories from the original variable (RES_010), which appears on the Restricted version only, are collapsed in order to maintain confidentiality.
RESPONSE OPTIONS FOR DVRES	1 = I was born in Canada 2 = 1 to 10 years 3 = 11 or more years 99 = Not Stated
COVERAGE FOR DVRES	Coverage: All Respondents
QUESTIONS	RES_010 (Question 6): *How many years have you lived in Canada?* 1 = I was born in Canada 2 = 1 to 2 years 3 = 3 to 5 years 4 = 6 to 10 years 5 = 11 or more years 99 = Not Stated Coverage: All Respondents *Note that RES_010 has been removed from the PUMF due to confidentiality reasons. Notes: • Code "I was born in Canada" for respondents who responded "I was born in Canada" to question 6 (RES_010 = 1). • Code "1 to 10 years" for respondents who responded either "1 to 2 years", "3 to 5 years" or "6 to 10 years" to question 6 (2 <= RES_010 <= 4). • Code "11 or more years" for respondents who responded "11 or more years" to question 6 (RES_010 = 5). • Code "Not Stated" for respondents who did not respond to question 6 (RES_010 = 99).

Derived Variable for Urban/Rural

The following describes the derived variable in the PUMF in classifying each record as being urban or rural.

DERIVED VARIABLE	DVURBAN
OBJECTIVE	To classify the respondent's school location as urban or rural.
RESPONSE OPTIONS	1 = Urban 2 = Rural
QUESTION	The school's postal code was used to calculate this variable.
COVERAGE	All Respondents Using the Statistical Area Classification system variable SACTYPE obtained from the Postal Code Conversion File plus Version 7A, respondents were assigned a rural or urban status based on the postal code of the school they attended at the time of their participation in the survey. The SACTYPE variable distinguishes among census metropolitan areas (all of which are tracted), tracted versus untracted census agglomerations, and the residual area not in any census metropolitan area or census agglomeration ("rural and small town Canada"), with the latter further classified by the relative importance of commuting flows to work in any census metropolitan area or census agglomeration (CMACA) --also known as "metropolitan influence zones" or MIZ. SACTYPE has the following categories: 1 = CENSUS METROPOLITAN AREA 2 = TRACTED CENSUS AGGLOMERATION 3 = NON-TRACTED CENSUS AGGLOMERATION 4 = NON-CMACA, STRONG CMACA INFLUENCE 5 = NON-CMACA, MODERATE CMACA INFLUENCE 6 = NON-CMACA, WEAK CMACA INFLUENCE 7 = NON-CMACA, NO CMACA INFLUENCE 8 = NON-CMACA, TERRITORIES
NOTES	• Categories 1 to 3 would be categorized as Urban (DVURBAN = 1) and categories 4 to 8 would be categorized as Rural (DVURBAN = 2).

- Respondents with school Postal Codes for which the program was unable to compute SACTYPE were imputed into either an Urban or Rural category by using an older version of the Postal Code Conversion File (Version 6D).

Derived Variable for Socioeconomic Status (SES)

The following describes the derived variable in the PUMF for socioeconomic status (SES).

DERIVED VARIABLE	DVHHINC2
OBJECTIVE	This variable reflects the median household income for the area where the respondent's school is located.
QUESTION	The school's postal code was used to calculate this variable.
COVERAGE	All Respondents

The forward sortation area (FSA), which is the first three digits of the postal code of the respondent's school, was merged onto the Canadian 2016 Census data on median household income that was sorted by FSA. The median household income was then rounded to the nearest $5000. Outliers were grouped so that at least 5 schools are associated with each income category. |

7.4   Weighting

Survey weights are needed to derive population estimates from the survey sample. In a simple random sample, every unit in the population has the same probability of being drawn. The *fraction* of the population that is sampled is the sample size divided by population size. To calculate the weight of each sampled member, one should multiply each member by $1/fraction$. If the sample size was 100 and the population was 100,000, then the weight of each sampled member would be 1,000. This means that any sampled member's response is taken to represent 1,000 identical responses in the population. Constant sampling fractions do not apply to complex survey designs such as those used in the CSTADS. In complex survey designs, the sample data must be multiplied by appropriate weights that reflect the different sampling fractions. This survey weight is included in the PUMF and must be used to derive meaningful population estimates from the survey. See Section 8 for guidelines on using the data.

The development of the survey weights was accomplished in a series of stages detailed below. In the first stage, a weight (W$_{1j}$) was created to account for the school selection within health region and school strata. A second weight (W$_{2jg}$) was calculated to adjust for student non-response (Stage 2). The product of these two weights (Stage 3) was re-based province by province and trimmed (Stage 4). The resulting re-based and trimmed weights were calibrated to the provincial sex and grade distribution so that the total of the survey weights by sex, grade and province would equal the actual enrolments in those groups for the target population (Stage 5). Finally, bootstrap weights (see Stage 6) were generated to attach to the data file.

Stage 1: Calculation of W$_{1j}$

Within each provincial sampling frame, there are two to four health region categories: "low" and "high" for all provinces except Prince Edward Island, Nova Scotia, and New Brunswick, and "urban" for Newfoundland and Labrador, Québec, Ontario, and Alberta, and for Nova Scotia only, stratification by specific health region. Two school-level categories ("elementary" and "secondary") are also defined. Crossing these stratifications yields eight strata in Nova Scotia, six strata in Newfoundland and Labrador, Ontario and Alberta, and four in Manitoba, Saskatchewan and British Columbia. In Prince Edward Island, there is only one health region in the province, resulting in two strata. In Québec, the provincial sampling frame contained only secondary schools, resulting in three strata. Within each stratum, in each province, schools were selected based on simple random sampling.

New Brunswick was sampled differently due to the collaboration with the NBSWS as described in Section 4.4. For the purpose of creating the weights, the whole province of New Brunswick will be considered as one stratum.

Within each stratum, in each province, schools were randomly selected. For school *j*, W$_{1j}$ has been computed as

$$\mathbf{W_{1j} = 1/\pi_{1j}}$$

where π_{1j} is the probability of inclusion at stage 1 for school j*, and where

$$\mathbf{\pi_{1j}} = \ell / \mathsf{L}$$

ℓ = Number of participating schools in the given stratum, and

L = Total number of schools in the stratum.

Stage 2: Calculation of W$_{2jg}$

Within each selected school we computed the response rate for the students by grade.

$$\pi_{2jg} = \frac{n_j(g)}{N_j(g)}$$

where

$n_j(g)$ is the number of participating students in grade g in school j

$N_j(g)$ is the total number of eligible students in grade g in school j

Then
$$W_{2jg} = 1/\pi_{2jg}$$

Stage 3: Un-Calibrated Final Weight

The final un-calibrated weight is based on the product of w$_{1j}$ and w$_{2jg}$.

$$w_{3jg} = \left(w_{1j} * w_{2jg} \right)$$

Stage 4: Re-basing and Trimming the Un-Calibrated Final Weight

The final un-calibrated weight is then re-based province by province using the formula:

$$rw_{3jg} = \frac{w_{3jg}}{\sum w_{3jg}} \quad * \text{(sample size within respondent's province)}$$

Where rw_{3jg} is the re-based weight and $\sum w_{3jg}$ is the sum of the weights within the respondent's province.

Consistent with standard practice, the re-based weight is then trimmed at the level of 10, such that all respondents with a re-based weight larger than or equal to 10 are equated to 10.

Stage 5: Calibration of Survey Weights

The re-based and trimmed weights described above were then calibrated using school administrative datasets that include the total student enrolment of the target population by grade (grades 7 through 12) for each province gathered from the sampling frame. Province, grade, and sex calibration were used to adjust the sampling weights so that estimated numbers of students in these domains reproduce

known population numbers exactly. The final weight variable was defined as WTPP in the 2018/2019 CSTADS PUMF. The calibration was done by dividing the final un-calibrated, re-based and trimmed weight by the t-ratio, as in this formula:

$$WTPP=\left(\frac{rw_{3\,jg}}{t-ratio}\right)$$

The t-ratio is created by dividing the weighted frequency counts for each province by grade by sex (created by the final un-calibrated, re-based and trimmed weight), by each province's population numbers by grade by sex. The provincial population numbers by grade are obtained from the sampling frame. We do not have sex information in our sampling frame, so we used the sex ratio, by age, from Statistics Canada's post-censal CANSIM tables for 2018 (formerly Table 051- 0001) where the age-based sex ratios were used for the following grades: 12 years = grade 7, 13 years = grade 8, 14 years = grade 9, 15 years = grade 10, 16 years = grade 11, and 17 years = grade 12.

Stage 6: Construction of Bootstrap Weights

Statisticians use resampling methods such as the bootstrap to more precisely estimate sampling error. Bootstrap weights for each province were constructed separately using the following process. The bootstrap weights can be accessed through Propel's Data Repository.

1) Within each stratum (health region smoking rate crossed with grade-level) of *n* participating schools, we resampled *n* schools with equal probability and with replacement (this is called "unrestricted random sampling" in the SAS procedure, PROC SURVEYSELECT). A bootstrap weight is created from a sample selected where some schools among the participating schools were not selected, and some were selected once and some were selected multiple times.

2) The weight was recalculated and adjusted for the resampling inference based on the method of Rao and Wu (1988)[7].

3) Finally, the new weight was recalibrated to the provincial enrolment figures using the administrative datasets.

Six thousand (6,000) such bootstrap samples were computed. The average of sets of twelve bootstrap weights were used to create a set of 500 averaged bootstrap weights.

The formula for the weight adjustment is obtained as follows:

Let
$$\lambda_{1i}=\sqrt{\frac{n_i}{n_i-1}\left(1-\frac{n_i}{N_i}\right)}$$

---

[7] Rao, J.N.K. & Wu, C.F.J. (1988). Resampling inference with complex survey data. Journal of the American Statistical Association 83, 231-241.

where N_i is the number of schools in a given stratum and n_i is the number of schools actually selected in that stratum.

Let w_{ij} be the smoothed calibrated main weight for student j in school i.

The bootstrap weight w_{ij*} would then be given by

$$w_{ij}[1 - \lambda_{1i} + \text{(number of times school } i \text{ has been resampled)}* (\lambda_{1i})].$$

7.5   Suppression of Confidential Information

It should be noted that not all variables are included in the PUMF, as actions have been taken to protect the anonymity of individual survey respondents. Specifically, the following variables have been removed from the PUMF to help ensure anonymity.

- school board & school identifiers (including postal code) with the exception of school identification number
- class identifiers
- strata identifiers
- age
- gender
- sexual orientation
- years lived in Canada (prior to collapsing categories)

8.0 Guidelines for Tabulation, Analysis and Release

This section details guidelines for users when tabulating, analyzing, and publishing or otherwise releasing any data derived from the survey data files. With the aid of these guidelines, users of the 2018/2019 CSTADS PUMF should be able to produce the same figures as those produced by any statistician and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines. Please note that these guidelines are consistent with past implementations and have been adapted from the 2002 YSS User Guide written by Statistics Canada.[8]

8.1   Survey Errors

The estimates derived from this survey are based on a sample of schools. Somewhat different estimates might be obtained if a complete census had been taken using the same questionnaire, data collection staff, and processing methods. The difference between the estimates obtained from the sample and those resulting from a complete count taken under similar conditions are called the sampling error of the estimate.

Errors that are not related to sampling may occur at almost every phase of survey implementation. Administrators may misunderstand instructions, respondents may make errors in answering questions,

---

[8] Statistics Canada (2002). Microdata User Guide: Youth Smoking Survey 2002. Accessible at: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4401&lang=en&db=imdb&adm=8&dis=2

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 included 1) the use of protocols that have been validated in previous studies of school-based data collections related to youth smoking; 2) detailed instructions for teachers and participating students; 3) extensive training of project staff with respect to survey procedures; 4) procedures to ensure that data capture errors were minimized; and 5) coding and editing quality checks to verify the processing logic.

8.2   Use of Survey Weights

Why should survey weights be used?

When producing simple population estimates, including the production of ordinary statistical tables, users must apply the proper sampling weights. There are two reasons why a survey weight variable should be used when performing analyses.

1) **Total population versus sample size[9]**. Users may want results based on population figures instead of estimates based on the sample of individuals included in the study. For example, the CSTADS survey weight, when used, will produce results based on a national population estimate of N, which represents all the students in the participating provinces (grades 7-12) instead of n, which is the total number of students who actually completed the survey. The latter is known as the sample size of the CSTADS.

2) **Adjusting for sampling method.** The second use of survey weights is to adjust for sampling methods. If every member of a population has an equal probability of being selected in a sample, each case would carry the same survey weight and the survey weight for all individuals would be 1. In reality, however, CSTADS sampling employed more complex considerations (described in Section 4.0). As a result, individuals did not have an equal probability of being selected. To correct for this unequal probability or chance of being selected, the survey weight variable was created. In short, using the survey weight variable permits the user to make generalizations to the population from which the sample was drawn.

If proper weights are not used, the estimates derived from the PUMF cannot be considered as representative of the survey population, and will not correspond to estimates produced by Health Canada.

8.3   Rounding Guide

Users are urged to adhere to the following guidelines regarding the rounding of such estimates:

1) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units

---

[9] Sample size refers to the number of records in the final data set.

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 1. 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.

2) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding un-rounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.
3) Averages, proportions, rates and percentages are to be computed from un-rounded components (i.e., numerators and/or denominators) and then are to be rounded 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.
4) Sums and differences of aggregates (or ratios) are to be derived from their corresponding un-rounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding.
5) Under no circumstances are un-rounded estimates to be published or otherwise released by users. Un-rounded estimates imply greater precision than actually exists.

8.4   Use of Weights for Variance Estimation

The weighting procedure in some statistical analysis software packages may not be appropriate for sample survey analysis. The user should understand how the weight variable is used within the software package.

The calculation of more precise variance estimates and coefficients of variation requires detailed knowledge of the design of the survey. Such details cannot be given in the PUMF since confidentiality must be respected.

A preferable method of calculating variances is available that takes account of the sample design by using the bootstrap weights which are provided as a separate data file. The correct variance for the mean bootstrap can be computed by specifying balanced repeated replication (BRR) with an appropriate Fay factor[10]. For the variance computation specify BRR with the Fay factor set to $\rho = 1 - \sqrt{1/12}$ = 0.71132487. The number "12" in the formula stems from each mean bootstrap sample representing an average over 12 bootstrap samples. The replicate weights commands can be used to specify the bootstrap weights. This procedure, which can be applied in statistical software such as STATA and SAS, creates reliable estimates of the variance for both simple estimates such as estimates of totals, proportions and ratios, and more complex analyses such as linear or logistic regression.

Health Canada employed STATA [11]for all analyses of the 2018/2019 CSTADS. The ***svyset*** commands are used to account for the complex survey design. The ***svyset*** statement to be used in STATA with the 2018/2019 CSTADS dataset would have the following form:

svyset [pweight=wtpp], brrweight(bsw1-bsw500) vce(brr) fay(0.71132487)

---

[10] Judkins, D. 1990. Fay's Method for Variance Estimation. Journal of Official Statistics, 6(3), 223-239
[11] StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP

Declaring **pweight=wtpp** tells STATA that the survey weight (which is often called the probability weight) is the variable **wtpp**. The option **vce(brr)** states that the variance estimation approach to use is BRR. The option **fay(0.71132487)** states that the BRR variance estimation approach is to use a Fay's adjustment of 0.71132487. The option **brrweight(bsw1-bsw500)** states that the names of the BRR weight variables are **bsw1**, **bsw2**, …, **bsw500**. This option can also be designated as **brrweight(bsw*)** provided there are no variables other than the bootstrap weight variables whose names begin with "bsw".

Finally, the **mse** option, which can be added to the end of the **svyset** statement, tells STATA to calculate the variance using squared differences between bootstrap estimates and the full- sample estimate of the quantities of interest. If this option is not included, STATA uses squared differences between each bootstrap estimate and the mean of all the bootstrap estimates. Both approaches are acceptable and should yield approximately the same result.

Equivalently in SAS, the survey procedures (SURVEYFREQ, SURVEYMEANS, SURVEYLOGISTIC, SURVEYREG, etc.) can be used to account for the complex survey design. This is done using the VARMETHOD=BRR (FAY = 0.71132487) option for the procedure in question and by using the REPWEIGHTS BSW1 -- BSW500 statement and the WEIGHT WTPP statement.

SURVEYREG produces some incorrect results when the REF=option is specified and BRR variance estimation is used. This is only an issue with certain releases of SAS. More information can be found on the SAS website (http://support.sas.com/kb/59/127.html). Please note, this website only mentions the Jackknife method, but it is in fact also an issue with BRR variance estimation.

The SAS survey procedures are equivalent to using the **mse** option in STATA. Another option is to use the Bootvar program available in SAS. It is made up of macros that compute variances for totals, differences between ratios and for linear and logistic regression.

Bootvar may be downloaded from Statistics Canada's Research Data Centre website (http://www.statcan.gc.ca/pub/12-002-x/2014001/article/11901-eng.htm#a8). Users must accept the Bootvar Click-Wrap License before they can read the files. There is a document on the site explaining how to adapt the system to meet users' needs.

8.5   Definitions of Types of Estimates: Categorical and Quantitative

Before discussing how the CSTADS 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 PUMF for CSTADS.

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 students who ever smoked a whole cigarette or the proportion of smokers who usually buy cigarettes from a friend or someone else are examples of such estimates. 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:

> Question: Have you ever smoked a whole cigarette?

Responses: Yes / No

Questions: Where do you usually get your cigarettes?

Responses: I do not smoke / I buy them myself at a store / I buy them from a First Nation Reserve (i.e., delivery service) / I buy them on a First Nation Reserve / I buy them from a friend / I buy them from someone else / I ask someone to buy them for me / etc.

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 $\widehat{X}/\widehat{Y}$ where $\widehat{X}$ is an estimate of surveyed population quantity total and $\widehat{Y}$ is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate in the 2018/2019 CSTADS is the average number of whole cigarettes smoked over the last 7 days among all current smokers. The numerator $\left(\widehat{X}\right)$ is an estimate of the total number of whole cigarettes smoked over the last 7 days and the denominator $\left(\widehat{Y}\right)$ is the number of persons who reported smoking over the last 7 days.

Tabulation of Categorical Estimates

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

a) summing the final weights of records having the characteristic of interest for the numerator $\left(\widehat{X}\right)$,

b) summing the final weights of records having the characteristic of interest for the denominator $\left(\widehat{Y}\right)$, then

c) dividing estimate a) by estimate b) $\left(\widehat{X}/\widehat{Y}\right)$.

Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the CSTADS PUMF 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 whole cigarettes smoked in the past 7 days prior to the survey by students in grade 9 (secondary III in Québec), multiply the value reported in the derived variable DVCIGWK (number of whole cigarettes smoked in the past 7 days prior to the survey) by the final weight for the record (WTPP), then sum this product for all records where DVCIGWK < 996 and GRADE equals 9.

8.6   Coefficient of Variation Release Guidelines

Before releasing and/or publishing any estimate from the 2018/2019 CSTADS, 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 Section 8.1. 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 (i.e., standard error divided by the estimate, multiplied by 100) as shown in the table below.

First, determine the number of respondents who contributed to the numerator in the calculation of the estimate. **If this number is less than 30, the weighted estimate must be considered to be of unacceptable quality and cannot be released.**

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 in Table 7. Apply these quality level guidelines to weighted rounded estimates. Unacceptable quality level estimates cannot be released and marginal level estimates can only be released with the warning to caution subsequent users outlined in Table 7.

Table 7: Quality Level Guidelines for Weighted Estimates

Quality Level of Estimate	Guidelines
Acceptable	Estimates have a sample size of 30 or more and low coefficients of variation in the range of 0.0% to 16.5%. No warning is required.
Marginal	Estimates have a sample size of 30 or more and high coefficients of variation in the range of 16.6% to 33.3%. Estimates should be flagged with the letter M (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.
Unacceptable	Estimates have a sample size of less than 30, or very high coefficients of variation in excess of 33.3%. It is not recommended to release estimates of unacceptable quality. Such estimates should be replaced with the letter U (or some similar identifier) and the following statement: **"Unreleasable due to low sample size."**

APPENDICES

APPENDIX A: 2018/2019 CANADIAN STUDENT TOBACCO, ALCOHOL AND DRUGS SURVEY QUESTIONNAIRE

APPENDIX B: COMPARABILITY OF QUESTIONNAIRE ITEMS ACROSS SURVEY CYCLES

Table B-1: CSTADS Questions in all cycles: 9 items

Question	Variable Name 2018/2019, 2016/2017 & 2014/2015 CSTADS	Variable Name 2012/2013 and previous Youth Smoking Survey (YSS)
Have you ever smoked 100 or more whole cigarettes in your life?	SS_040	SHUND0A1
Have you ever tried to quit smoking cigarettes?	SC_010	SEVRQTA1
Have you ever tried cigarette smoking, even just a few puffs?	SS_010	SPUFF0A1
How old are you today?	AGE	AGE
What was your sex at birth?⌀	SEX	SEX
What grade are you in?	GRADE	GRADE
On how many of the last 30 days did you smoke one or more cigarettes?	SS_060	SLST30A1
Have you ever smoked a whole cigarette?	SS_030	SWHOLEA1

⌀ Prior to the 2018/2019 questionnaire this question read: Are you…? Female Male

Table B-2: New Questions for the CSTADS 2018/2019 Cycle: 23 items

New Questions	Variable Name
What is your gender? *Gender refers to current gender which may be different from sex assigned at birth and may be different from what is indicated on legal documents.*	GENDER
Which of the following best describes you?	ORIENTATION
How many years have you lived in Canada?	RES_010
At any time during the next 12 months do you think you will smoke a cigarette?	TS_011
At any time during the next 12 months do you think you will use an e-cigarette (vape, vape pen, tank & mod)?	TV_010

Thinking back over the last 7 days, how many whole cigarettes did you smoke each day?
    Sunday
    Monday
    Tuesday
    Wednesday
    Thursday
    Friday
    Saturday
	
	WP_040a
	WP_040b
	WP_040c
	WP_040d
	WP_040e
	WP_040f
	WP_040g

Where do you usually get your cigarettes?	CA_011
Thinking about the last time you bought cigarettes in the last 12 months, how much did you pay?	PP_021
On how many of the last 30 days did you use an e-cigarette (vape, vape pen, tank & mod, e-juice)?	ELC_024

In the last 30 days, did you use any of the following?
 	
  a) Cigars, little cigars or cigarillos (plain or flavoured)
  b) Smokeless tobacco (chewing , pinch, dip, snuff, or snus)
  c) Nicotine patches, nicotine gum, nicotine lozenges, nicotine inhalers, or nicotine spray
  d) A water-pipe (hookah) to smoke shisha (tobacco)
  e) Heated tobacco products (iQOSTM or GloTM)
  f) E-cigarettes (vape, vape pen, tank & mod) with nicotine
  g) E-cigarettes (vape, vape pen, tank & mod) without nicotine
	
	TP_016
	TP_046
	TP_056
	
	TP_066
	TP_086
	ELC_026a
	ELC_026b

In the last 12 months, did you drink any of the following? e) Sweetened beverages with high alcohol content (7% or higher), (such as Four Loko, FCKD UP, Clubtails)	 NRG_050
In the last 12 months how did you usually get the marijuana or cannabis you used?	CAN_121
The use of cannabis was made legal for adults in Canada. Has it been easier to get marijuana or cannabis for yourself after legalization?	CAN_130
In the last 12 months, how often did you have alcohol AND marijuana or cannabis on the same occasion? (e.g., at a party, in the same evening, etc.)	CAN_140

Have you ever used a drug or substance to get high that was not what you thought it was?	UND_020

How much do you think people risk harming themselves when they do each of the following activities?
    e)   Use an e-cigarette with nicotine once in a while
    f)   Use an e-cigarette with nicotine on a regular basis
    g)   Use an e-cigarette without nicotine once in a while
    h)   Use an e-cigarette without nicotine on a regular basis
	
	PH_051
	PH_061
	PH_052
	PH_062

How difficult or easy do you think it would be for you to get each of the following types of substances, if you wanted some?
    b)   An e-cigarette with nicotine
    c)   An e-cigarette without nicotine
	
	ELC_041
	ELC_042

Which behaviours are allowed, or do you think are allowed, at your house?
    a)   smoking cigarettes?
    b)   smoking cannabis?
    c)   vaping e-cigarettes?
    d)   vaping cannabis?
	
	BEH_010
	BEH_020
	BEH_030
	BEH_040

On a typical school night (Sunday to Thursday), what time do you…
    a)   go to bed? – hour –minute
    b)   fall asleep? –hour –minute
	
	SLP_020 & SLP_021
	SLP_030 & SLP_031

On a typical school morning (Monday to Friday), what time do you…
    a)   wake up? – hour –minute
    b)   leave your house to get to school? – hour –minute
    c)   arrive at school? – hour –minute
	
	SLP_040 & SLP_040
	SLP_050 & SLP_051
	SLP_060 & SLP_061

On a typical WEEKEND, what time do you…
    a)   go to bed? (Friday and Saturday night) – hour –minute
    b)   fall asleep? (Friday and Saturday night) – hour –minute
    c)   wake up? (Saturday and Sunday morning) – hour –minute
	
	SLP_070 & SLP_071
	SLP_080 & SLP_081
	SLP_090 & SLP_091

Do you use electronics (e.g., TV, video games, computer, tablet, or smartphone) before bedtime?	SLP_100

In the last two weeks, how often have you…
    a)   felt satisfied with your sleep?
    b)   arrived late to class because you overslept?
    c)   fallen asleep in a morning class?
    d)   fallen asleep in an afternoon class?
    e)   stayed up until at least 3am?
    f)   stayed up all night?
    g)   slept past noon?
    h)   felt tired, dragged out, or sleepy during the day?
    i)   needed more than one reminder to get up in the morning?
    j)   had an extremely hard time falling asleep?
    k)   gone to bed because you just couldn't stay awake any longer?
    l)   struggled to stay awake while reading, studying, or doing homework?
	
	SLP_110
	SLP_120
	SLP_130
	SLP_140
	SLP_150
	SLP_160
	SLP_170
	SLP_180
	SLP_190
	SLP_200
	SLP_210
	SLP_220