ITC Four Country Smoking and Vaping Survey Wave 5 (4CV5, 2024) Technical Report

Version Date: 13 November 2025

Funding: The ITC Four Country Smoking and Vaping Survey (also termed the 'Project 1 Nicotine Product Adult Survey') was supported by grants from the US National Cancer Institute (P01 CA200512), the Canadian Institutes of Health Research (FDN-148477), Health Canada (4CV5), and the National Health and Medical Research Council of Australia (GTN2019252).

Ethics Clearance: The survey protocols and all materials of Wave 4 ITC Four Country Smoking and Vaping Survey, including the survey questionnaires, were cleared for ethics by Office of Research Ethics, University of Waterloo, Canada (REB#20803/30570, REB#21609/30878); Research Ethics Office, King's College London, UK (IRB RESCM-17/18-2240), Human Ethics, Research Management Office, University of Queensland, Australia (2016000330/HREC1603, IRB#2022/HE001187); and Institutional Review Board Medical University of South Carolina (waived due to minimal risk). All participants provided consent to participate.

Suggested Citation: ITC Project. (2025, November). *ITC Four Country Smoking and Vaping Survey, Wave 5 (4CV5, 2024) Technical Report*. University of Waterloo, Waterloo, Ontario, Canada; Medical University of South Carolina, Charleston, South Carolina, United States; the University of Queensland, Australia; King's College London, London, United Kingdom.

Table of Contents

Acknowledgements

The Wave 5 ITC Four Country Smoking and Vaping (4CV) Survey Technical Report was prepared from documents and other material that the International Tobacco Control Policy Evaluation Project (ITC) Research Team has created since 2015. The compilation of these documents and preparation of the material that accompanies those documents was conducted by Geoffrey T. Fong, Principal Investigator of the ITC Project; Mary Thompson, Co-Investigator and Director of the Data Management and Analysis Team; Christian Boudreau, Associate Director of the Data Management and Analysis Team; Pete Driezen, Senior Analyst; Grace Li, Data Analyst; and ITC Project Management members Thomas Agar, Anne C.K. Quah, and Janine Ouimet. We thank the Co-Investigators from the participating sites across the four countries (University of Waterloo, Canada; Medical University of South Carolina and Roswell Park Comprehensive Cancer Centre, US; King's College London, UK; Cancer Council Victoria and the University of Melbourne, Melbourne, Australia; and University of Queensland, Australia), Co-Investigators from other institutions (University of South Carolina, Georgetown University, Virginia Tech, Ohio State University, Governors State University, Deakin University), and the survey firms (Roy Morgan for Australia; Leger for Canada; Ipsos UK (previously Ipsos MORI) and Rakuten Insight for England; and Ipsos Public Affairs for the United States); and the Statistical Consulting and Survey Research Unit at the University of Waterloo, Canada (for hosting the survey).

We thank our funders for their support:

U.S. National Cancer Institute (United States)
Health Canada and Canadian Institutes of Health Research (Canada)
National Health and Medical Research Council (Australia)

ITC Four Country Smoking and Vaping Project Team

Canada (University of Waterloo)

Geoffrey T. Fong, Ph.D. (Principal Investigator)
Shannon Gravely, Ph.D. (Deputy Principal Investigator)
Christian Boudreau, Ph.D.
David Hammond, Ph.D.
Mary E. Thompson, Ph.D.
Pete Driezen, Ph.D.
Anne C.K. Quah, Ph.D.

Project and Survey Management Support
Thomas Agar, M.A.
Janine Ouimet, M.Sc.

Data Analysis Support
Grace Li, M.Math.
Gang Meng, Ph.D.

Christopher Bryan Lok, Ph.D.
Jamie Wheeler, B.Math.
Edward Bauer, B.Sc.
Simon Thompson, B.A.
Dana Komer, B.A.

England

Ann McNeill, Ph.D. (Co-Principal Investigator), King's College London
Katherine A. East, Ph.D., (Principal Investigator), Brighton and Sussex Medical School, King's College London

United States

Andrew Hyland, Ph.D., (Principal Investigator), Roswell Park Comprehensive Cancer Center
Maansi Travers, Ph.D., (Deputy Principal Investigator), Roswell Park Comprehensive Cancer Center
K. Michael Cummings, Ph.D., MPH, Medical University of South Carolina
Richard J. O'Connor, Ph.D., Roswell Park Comprehensive Cancer Center
Ce Shang, Ph.D., Ohio State University
James T. Thrasher, Ph.D., University of South Carolina
David T. Levy, Ph.D., Georgetown University
Maciej L. Goniewicz, Ph.D., Roswell Park Comprehensive Cancer Center
Tracy Smith, Ph.D., Medical University of South Carolina
Kai-Wen Cheng, Ph.D., Governors State University

Australia

Coral Gartner, Ph.D., (Principal Investigator), The University of Queensland
Kylie Morphett, Ph. D., The University of Queensland
Gary K.C. Chan, Ph. D., The University of Queensland
Ron Borland, Ph.D., Deakin University
Hua Yong, Ph.D., Deakin University

Abbreviations and Terminology

The abbreviations and terms listed here are used for this report. For academic publications, see *Appendix D – Terminology guidelines for academic publications*.

Abbreviation	Definition
ITC	International Tobacco Control Policy Evaluation Project
4C	ITC Four Country Survey
4CV	ITC Four Country Smoking and Vaping Survey
ETM	Experimental Tobacco Marketplace
CCV	Cancer Council Victoria
Dynata (prev SSI)	Dynata was a panel partner for RM in AU, and for Ipsos UK and Rakuten in England. Prior to 4CV3, Dynata was SSI (Survey Sampling International)
DMC	Data Management Centre
Ipsos	A survey firm recruiting part of the ITC sample in the United States. Previously known as GfK Knowledge Panel.
Ipsos UK	Previously named Ipsos Mori. A survey firm recruiting part of the ITC sample in England and part of the sample in the United States
Leger	Survey firm recruiting the ITC sample in Canada
TF-ONP	Tobacco-free oral nicotine products. An emerging class of nicotine product (often lozenges, pouches, etc) that delivers nicotine to users. Contains no tobacco.
PMG	Project Management Group
Rakuten	Rakuten Insight. Survey firm recruiting part of the ITC sample in England
RM	Roy Morgan. Survey firm recruiting the ITC sample in AU using their proprietary panel, Single Source.
SMG	Survey Management Group
SCSRU	Survey Research Centre. Survey firm programming and hosting the ITC 4CV Survey. Based at the University of Waterloo, Canada.
SSI (now Dynata)	Survey Sampling International. At 4CV3, SSI changed to **Dynata**. Dynata was a panel partner for RM in AU, and for Ipsos UK and Rakuten in England.
UoM	University of Melbourne
UQ	The University of Queensland
UW	University of Waterloo
VT	Virginia Tech
NVP	Nicotine Vaping product
HTP	Heated tobacco product
PIN	Personal Identification Number
RDD	Random digit dialing

Terminology	Definition
Heated tobacco product (HTP)	Product that heats actual tobacco in the form of sticks or capsule, or loose tobacco. Some HTP may also contain liquid, but what matters is that they contain actual tobacco.
Nicotine vaping product (NVP)	A vaping product (sometimes called an e-cigarette) that heats a liquid only. It does not contain actual tobacco. The liquid often contains nicotine and is often flavoured.
Smoker:	Someone who currently smokes at least occasionally, and has smoked at least 100 cigarettes, irrespective of NVP use.
Vaper:	Someone who currently vapes daily or weekly, irrespective of cigarette smoking
Long-term former smoker*:	Someone who has quit smoking for more than 24 months and less than 5 years, has smoked at least monthly in the past, and has smoked at least 100 cigarettes; **OR** has quit for more than 5 years but currently vapes at least weekly, or was vaping at least weekly at 4CV1. **See Table 2 below for differences in definition and inclusion across waves and countries.* *Note: Terms "former smoker" and "quitter" are equivalent.*
Recent former smoker:	Someone who has quit for less than 24 months, has smoked at least monthly in past, and has smoked at least 100 cigarettes, irrespective of NVP use. *Note: Terms "former smoker" and "quitter" are equivalent.*
Recontact:	Describes a respondent who participated in the previous wave and was re-invited to participate at this wave, or any survey form or activities customized for that group.
Replenishment:	Describes a new respondent who was recruited at this wave, or any survey form or activities customized for that group.
Panelist:	Any respondent from an online panel, either from a survey firm or a survey firm's partner.
ITC-owned respondent:	A respondent for whom ITC has access to his/her contact information and can invite him/her directly. Recruitment sources could be from: a) *a telephone RDD sampling frame created for the ITC during 4C (i.e., ITC-owned 4C cohort), or* b) *an opt-in panel partner firm in the US during 4C9 or 4CV1 (i.e., new ITC-owned).*
ITC-owned 4C cohort respondent:	A respondent who was recruited *from a telephone RDD sampling frame* at 4C (i.e., any wave prior to 4CV1).
New ITC-owned respondent:	A respondent who was recruited from a US opt-in panel partner at 4CV1 or 4CV2.
Panel-owned respondent:	A respondent who belongs to a web panel survey firm or panel partner. ITC does not have access to their contact information and all communications must take place through the relevant survey and/or partner panel firm.
Complete:	Respondent who submitted a completed survey.
Incomplete:	Respondent who was invited but did not access the survey.
Partial:	Respondent who accessed the survey but did not submit a completed survey.
Survey firm:	A country survey firm who recontacted previous wave respondents and/or recruited new respondents from their own (or a partner) online panel and/or programmed some or all of the survey.
Panel partner firm:	A separate survey firm working in collaboration with the country survey firm to recruit recontact and/or replenishment respondents from an alternate online panel.
Panel ID:	A survey firm's unique specific alphanumeric identification code.

Personal Identification Number (PIN):	A 7-digit alphanumeric code provided by SCSRU to link a respondent to his/her web survey record.
UniqID/unique ID	Unique identifier assigned by ITC to a newly recruited respondent who has completed their first survey; the respondent retains this same identifier when recontacted in subsequent waves.
Core data	Refers only to 4CV1. See 4CV3 Technical Report.
Full data	See definition (Table 2).
Main data	See definition (Table 2)

Preface

This report describes procedures used during the fifth wave of the International Tobacco Control Four Country Smoking and Vaping Survey (4CV5). 4CV5 Survey fieldwork was conducted approximately 2 years after Wave 4 (Table 1).

Table 1. 4CV Fieldwork dates

4CV Wave	Survey fieldwork dates	
	Start date	End date
4CV1	July 7, 2016	November 22, 2016
4CV2	February 23, 2018	July 9, 2018
4CV3	February 24, 2020	June 1, 2020
4CV4	August 4, 2022	December 17, 2022
4CV5	**July 17, 2024**	**February 21, 2025***

*Only a very small number of respondents were recruited in 2025.

Substantial changes at Wave 5 compared to Wave 4

- **The Experimental Tobacco Marketplace (ETM)** Sub-study was fielded at Wave 5, after a hiatus at Wave 4.

Format of this report

A **terminology** list for important terms and abbreviations used in this report.

Section 1 introduces this project and its **objectives**.

Section 2 provides **methods statements** and **ethics statements** for papers using 4CV4 data.

Section 3 describes the **sample size** and **eligibility criteria** (for cohort and replenishment samples).

Section 4 provides **survey questionnaire content** and the content development process.

Section 5 describes the **survey firms (panels)**.

Section 6 provides the details of the **study procedures**.

Section 7 reports on the sample **targets** and **outcome rates** (e.g., retention rate; response rate; cooperation rate). It also includes survey length values.

Section 8 describes the **ANDS Image Upload Sub-study** associated with 4CV4.

Section 9 reports on **quality control and uniformity** procedures applied to the sample, and **retention rates** from the previous wave.

Appendices A to G provide additional details about (A) sample size and cohort retention, (B) replenishment sample targets, (C) fieldwork challenges, (D) terminology guidelines for publications, (E) recruitment materials that used ITC phrasing (vs. panel standard phrasing), (F) sampling weights, design, and national benchmarks.

1 Introduction

Since 2002, the International Tobacco Control Policy Evaluation (ITC) Project has provided empirical data to inform governments and stakeholders about the effectiveness of public health policies designed to reduce the health, economic, and societal costs of tobacco use throughout the world. Established to evaluate key health policies implemented under the Framework Convention on Tobacco Control (FCTC)—the first-ever international public health treaty, adopted in May 2003 by all 192 member states of the World Health Organization —the ITC Project conducts representative cohort surveys in 29 countries.

From 2002 to 2015, the flagship Four Country (4C) Survey conducted nationally representative cohort surveys of adult smokers (n =1,500 to 2,000) in each of Australia (AU), Canada (CA), the United Kingdom (UK), and the United States (US). In 2016, the 4C Survey was expanded into the **ITC Four Country Smoking and Vaping (4CV) Survey,** to investigate the relationship between emerging alternative nicotine-delivery (ANDs) products and tobacco use, and to inform emerging policies related to nicotine use. The **4CV Survey** conducts concurrent longitudinal surveys of adult smokers and former smokers, NVP users (recruited since 2016), and heated tobacco product (HTP; recruited since 2020), snus, or tobacco-free oral nicotine users (both recruited since 2022) users in each of Australia, Canada, England, and the United States. These four countries have similar cigarette-smoking rates, but divergent policies on nicotine vaping product (NVP) use. While NVPs and alternative nicotine products may seem much less harmful compared to combustible tobacco (the known leading preventable cause of premature death in high-income countries), the effect of alternative nicotine product use on tobacco cessation, uptake, and sustained use is currently not well understood. Tobacco control experts are divided on whether policies should support, restrict, or ban NVP (and other ANDs products) use.

Over the past decade, demand for NVPs has grown rapidly, and data are required to inform public health policies. While NVPs seem less harmful compared to combustible tobacco products – the known leading preventable cause of premature death in high-income countries – the effect of NVPs use on tobacco cessation, uptake, and/or sustained use is currently not well understood. Tobacco control experts are still divided on whether policies should support, restrict, or ban NVP use.

The specific objectives of the ITC 4CV Survey Waves 4-6 are to:

1) Describe the complex interplay between cigarettes and NVPs/HTPs across the 4 countries plus Japan, Korea, and New Zealand, and to identify factors related to those patterns;
2) Evaluate the impact of policies on cigarettes (e.g., FCTC policies) and on NVPs/HTPs to determine: a) whether a policy has an effect (outcome evaluation), how and why the policy was effective (mediation), and c) for whom or under what conditions the policy was effective (moderation);
3) Evaluate the impact of policies among adults compared to among youth and young adults in Canada, England, and the US; and
4) Explore the interaction between smoking, using ANDs, and using cannabis; and the impact of COVID on smoking and using ANDs.

This report provides the methodological background and key statistical indicators for the ITC 4CV5 Survey and provides information on the sampling methods, procedures, and survey administration. This report also provides survey outcome rates, measures of representativeness, and guidelines for data analysis for 4CV5 data. Refer to the ITC Project website (https://itcproject.org/) for country-specific timelines of tobacco control and NVP policies in Australia, Canada, England, and the United States.

2 Methods Statements and Ethics Clearance

Below are method statements for a single wave and for both waves, followed by ethics statements.

2.1 Methods statement for using 4CV5 data only

For earlier waves, see the 4CV4 Technical Report available on the www.itcproject.org website.

Methodological details for each country are available via the ITC website (https://itcproject.org/methods). In brief, the ITC Four Country Smoking and Vaping Wave 5 (4CV5) Survey sample comprised the following subsamples: (1) recontact smokers and former smokers who had participated in the 4CV4 Survey, (2) newly recruited current smokers and former smokers (i.e., quit smoking in the previous 24 months) from country-specific panels, irrespective of vaping status, (3) recontact NVP users who had participated in the 4CV4 Survey, and (4) newly-recruited current NVP users (use a vaping device at least weekly) OR heated tobacco product users (use HTP at least weekly) OR snus users (use snus at least-weekly), OR 'tobacco-free oral nicotine product (TF-ONP) users' (use tobacco-free oral nicotine at least weekly) from country-specific panels, irrespective of smoking status. The newly-recruited smoker and NVP/HTP/snus/TF-ONP user samples in each country were designed to be representative of smokers and at-least-weekly NVP/HTP/snus/TF-ONP users respectively, and used either probability-based sampling frames or non-probability opt-in sampling frames, or a combination of these methods.

2.2 Methods for using 4CV1-5 data

Methodological details for each country are available via the ITC website (https://itcproject.org/methods). In brief, the ITC Four Country Smoking and Vaping Wave 1 (4CV1) Survey sample comprised the following cohorts: (1.1) recontact smokers and former smokers who participated in the previous waves of the ITC 4C Project, irrespective of vaping status, (1.2) newly recruited current smokers and former smokers (quit in the previous 24 months) from country-specific panels, irrespective of vaping status, and (1.3) newly recruited current NVP users (using a vaping device at least weekly) from country-specific panels. The smoker sample in each country was designed to be representative of smokers, and used either probability-based sampling frames or non-probability opt-in sampling frames, or a combination of these.

In the ITC Four Country Smoking and Vaping Wave 2-5 Surveys, most of respondents from the previous waves were invited to participate. Those who did not participate were replaced by newly recruited respondents from the same or similar sources, to maintain approximately the same sample sizes in the following subsamples: (2.1) smokers and former smokers (who had quit smoking in the previous 5 years), irrespective of vaping status; (2.2) current NVP users (using a vaping device at least weekly). At Wave 3, the sample was extended to include at-least-weekly NVP/HTP users; and at Wave 4, the sample was extended to include at-least-weekly NVP/HTP/snus/tobacco-free oral nicotine (TF-ONP) users.

2.3 Ethics Clearance Statement for 4CV5

For earlier waves, see the 4CV4 Technical Report available on the www.itcproject.org website.

The survey protocols and all materials of Wave 5 of the ITC Four Country Smoking and Vaping Survey, including the survey questionnaires, were reviewed and received clearance from research ethics committees at the following institutions: University of Waterloo, Canada (REB#20803/30570, REB#21609/30878); King's College London, UK (IRB RESCM-17/18-2240); Cancer Council Victoria, Australia (IRB HREC1603), and University of Queensland, Australia (IRB 2016000330/HREC1603, IRB#2022/HE001187). Research ethics committee review was waived at the Medical University of South Carolina. All participants provided informed consent to participate.

Table 2: Datasets available for 4CV1 to 4CV5

Name	Description	When to use
For previous waves please see the corresponding technical report posted on the www.itcproject.org website.		
4CV1 Core Sample	Subsample of the main sample created to unify the former smokers criteria across the 4 countries. Thus, records for respondents from CA, EN, and US who quit over 2 years ago (rather than 5 years) have been removed from the main sample so that the former smoker samples in each country meet the same definition as the eligibility criteria used in AU at 4CV1 (i.e., quit no more than 2 years ago).	*The core sample should be used to ensure matching eligibility criteria across AU and the other 3 countries prior to analyzing former smoker samples.*
4CV1 Main Sample	= Final 4CV1 sample with data from these groups removed: • speeders/satisficers, • long-term former smokers* ○ AU – all those quit over 2 years ago ○ CA, EN, US – only those quit over 5 years ago, • AU dedicated vapers sample. **Note on former smoker subsamples in the main sample at 4CV1, which vary based on countries:* • *In AU, it includes only respondents who had quit less than 2 years ago.* • *In CA, EN, and US, it includes both respondents who had quit less than 2 years ago, as well as those quit between 2 and 5 years.* *Thus, the main sample of AU former smokers (quit less than 2 years ago, thus all considered "recent") are defined differently than the main sample of CA, EN, US former smokers (which includes both "recent"= quit less than 2 years ago, and "long-term"= quit between 2 and 5 years ago). The core sample may need to be used when analyzing former smokers' data.*	*The main sample is the one intended to be used for the vast majority of analyses.*
4CV2 Main Sample	Part of the full valid data, with data from these respondents removed: • all AU dedicated vapers recruited from 4CV1 or 4CV2, • 5 respondents from CA (3 recontact and 2 replenishment) and 2 recontact respondents from the US. ○ all 7 respondents are long-term quitters and not vapers ○ 2 replenishment respondents do not have any weights 5 recontact respondents have longitudinal weights, but no cross-sectional weights.	*The main sample is the one intended to be used for the vast majority of analyses.*

Name	Description	When to use
4CV2 Full Sample	This sample includes everybody who completed the survey, with data from these groups removed: • 148 speeders/satisficers • 188 respondents with smoking or other status conflict between waves • 9 respondents who do not qualify through being smokers, former smokers, or vapers • 5 recontact respondents who were recruited as replenishment because survey firm(s) could not find their original 4CV1 PIN/ID	*The full data/sample was created to allow users to:* *conduct analyses using all vapers in AU (i.e., to include respondents recruited via the AU dedicated vapers sample)* *conduct longitudinal analyses that incorporates all respondents for whom a longitudinal weight was computed at 4CV2, that is all respondents who completed the 4CV1 Survey, and who were successfully recontacted at 4CV2. The full data/sample allows for analyses including the 5 recontact respondents who have the longitudinal weights.*
4CV3 Main Sample	Part of the full valid data, with data from these respondents removed: • all AU dedicated vapers recruited from 4CV1, 4CV2, or 4CV3 • 77 respondents quit smoking more than 5 years ago, and do not use e-cig/HTP (EN=18 ; CA=38; US=21) • 64 AU respondents quit smoking more than 2 years ago, and do not use e-cig/HTP	
4CV3 Full Sample	This sample includes everybody who completed the survey, with data from these groups removed: • 90 speeders/satisficers • 197 respondents with smoking or other status conflict between waves • 48 former EN respondents moved out of England • 212 respondents who do not qualify for recruitment requirement were wrongly recruited. Please note a respondent is counted more than once if he/she belongs to 2 or more groups	

Name	Description	When to use
4CV4 Main Sample	This sample includes everybody who completed the survey, with data from these groups removed: • 71 speeders (CA=24, US=21, EN=21 and AU=5) • 210 respondents (CA:64, US:77, EN:45, AU:24) who have quit smoking over 5 years and are not using any products (HTP, NVP, snus, or TF-OP) currently. • 1 EN respondent is out of survey frame • 11 respondents (CA=5 and US=6) NVP status claimed in the Wave 4 conflicts to the status reported in the previous waves. • 17 EN respondents are considered as fraudulent.	*The main sample is the one intended to be used for the vast majority of analyses.*
4CV4 Full Sample	This sample includes everybody who completed the survey, with data from these groups removed: • 71 speeders (CA=24, US=21, EN=21 and AU=5) • 1 EN respondent is out of survey frame • 11 respondents (CA=5 and US=6) NVP status claimed in the Wave 4 conflicts to the status reported in the previous waves • 17 EN respondents are considered as fraudulent.	*To conduct longitudinal analyses incorporating all respondents for whom a longitudinal weight was computed at 4CV4 (i.e., all respondents who completed 4CV3, and who were successfully recontacted at 4CV4).*
4CV5 Main Sample	This sample includes everybody who completed the survey, with data from these groups removed: • 71 speeders (CA=27, US=25, EN=12, and AU=7) • 4 respondents (CA=1,US=3) NVP status claimed in the Wave 5 conflicts to the status reported in the previous waves • 2 respondents (US) ONP status claimed in the Wave 5 conflicts to the status reported in the previous waves • 1 respondent (CA) didn't provide EC309v information • 4 respondents (US=2, EN=2) didn't hit the submit key. • 159 respondents(CA:58, US:75, EN:12,AU:14) who have quit smoking over 5 years and are not using any products (HTP, NVP, snus, or TF OP) currently. •	*The main sample is the one intended to be used for the vast majority of analyses.*
4CV5 Full Sample	This sample includes everybody who completed the survey, with data from these groups removed: • 71 speeders (CA=27, US=25, EN=12, and AU=7) • 4 respondents (CA=1, US=3) NVP status claimed in the Wave 5 conflicts to the status reported in the previous waves	*To conduct longitudinal analyses incorporating all respondents for whom a longitudinal weight was computed at 4CV4 (i.e., all respondents who*

Name	Description	When to use
	• 2 respondents (US) ONP status claimed in the Wave 5 conflicts to the status reported in the previous waves • 1 respondent (CA) didn't provide EC309v information • 4 respondents (US=2, EN=2) didn't hit the submit key.	*completed 4CV3, and who were successfully recontacted at 4CV4).*

3 SAMPLE SIZE AND ELIGIBILITY CRITERIA

3.1 Sample Size

Table 3: 4CV5 sample size (realized, with satisficers removed) by user type, by wave, and by country. Sample sizes for 4CV4 through 4CV2 are also provided for reference, with differences underlined.

Wave 5 Sample Size								
Sample groups	**Canada**		**England**		**US**		**Australia**	
User Type	**Cohort**	**Replenish-ment**	**Cohort** (Ipsos UK + Rakuten)	**Replenish-ment** (Rakuten)	**Cohort**	**Replenish-ment**	**Cohort**	**Replenish-ment**
ITC-owned	114	n/a	88	n/a	50	n/a	45	n/a
18+ Smokers/Recent former smokers	907	611	301	1406	1104	474	667	783
18+ NVP/HTP/snus/TF-ONP users (supplementary)	222	573	53	439	361	180		
Sub-Total	**1243**	**1184**	**442**	**1845**	**1515**	**654**	**712**	**783**
Total	**2427**		**2287**		**2169**		**1495**	

Wave 4 Sample Size								
Sample groups	**Canada**		**England**		**US**		**Australia**	
User Type	**Cohort**	**Replenish-ment**	**Cohort** (Ipsos UK + Rakuten)	**Replenish-ment** (Rakuten)	**Cohort**	**Replenish-ment**	**Cohort**	**Replenish-ment**
ITC-owned	159	n/a	109	n/a	66	n/a	82	n/a
18+ Smokers/Recent former smokers	1039	570	784	872	940	668	686	753
18+ NVP/HTP/snus/TF-ONP users (supplementary)	259	157	162	385	264	246		
Sub-Total	**1457**	**727**	**1055**	**1257**	**1270**	**914**	**768**	**753**
Total	**2184**		**2312**		**2184**		**1521**	
Wave 3 Sample Size								
Sample groups	**Canada**		**England**		**US**		**Australia**	
User Type	**Cohort**	**Replenish-ment**	**Cohort** (Ipsos UK + Rakuten)	**Replenish-ment** (Rakuten)	**Cohort**	**Replenish-ment**	**Cohort**	**Replenish-ment**
ITC-owned	224	n/a	135	n/a	94+61	n/a	137	n/a
25+ Smokers/Recent former smokers	1036	826	1156	1087	935	468	738	616
18-24 Smokers/Recent former smokers	104	641	23	996	55	434		
18+ NVP/HTP (supplementary)	297	560	272	231	213+6	252		
Sub-Total	**1661**	**2027**	**1586**	**2314**	**1364**	**1164**	**875**	**616**
Total	**3688**		**3900**		**2528**		**1491**	
Wave 2 Sample Size								
Sample groups	**Canada**		**England**		**US**		**Australia**	

User Type	Cohort	Replenish-ment	Cohort (Ipsos UK + Rakuten)	Replenish-ment (Rakuten)	Cohort	Replenish-ment	Cohort	Replenish-ment
ITC-owned	331	n/a	204	n/a	188	n/a	199	n/a
25+ Smokers/Recent former smokers	1025	1068	1179	1810	892	696		
18-24 Smokers/Recent former smokers	148	515	2+90	984	48	516		
18+ NVP (supplementary)	364	332	256	323	77	431		
Sub-Total	**1868**	**1915**	**1731**	**3117**	**1205**	**1643**	**853**	**662**
Total	**3783**		**4848**		**2848**		**1515**	

Table 4. Sources and sample sizes for 4CV1 respondents.

Country, subsample	Sources	Main* Sample size	Core* sample
AU, smokers/former smokers 18+	ITC 4C cohort (retention rate 43.1%)	515	515
	Roy Morgan Single Source (probability-based)	504	504
	Survey Sampling International	485	485
CA, smokers/former smokers 18+	ITC 4C cohort (retention rate 41.6%)	567	515
	Léger Marketing, where possible from their probability-based panel	2439	2439
CA, additional vapers 18+	Léger Marketing	727	727
EN, smokers/former smokers 18+	ITC 4C cohort (retention rate 35.7%)	254	232
	Ipsos	3519	3519
EN, additional vapers 18+	Ipsos	551	551
US, smokers/former smokers aged 25+	ITC 4C cohort	1372	1296
	GfK Knowledge Panel (probability based)	127	127
	Ipsos	212	212
US, smokers/formers smokers aged 18-24	ITC 4C cohort (retention rate 44.2%)	6	6
	Lucid (GfK partner panel, opt-in)	496	496
	Ipsos	26	26
US, additional vapers 18+	Ipsos	494	494
Total		12294	12144

*Main sample (see Section 6) = Final realized 4CV1 sample with speeders/satisficers removed, long-term former smokers (quit >5y ago) removed, and the AU additional vapers sample removed. *Note on former smoker subsamples in the main sample: In each of CA, EN, and US, cohort respondents who had quit >2y ago were interviewed, while in AU, only respondents who had quit <2y ago were interviewed. Thus, the main sample AU former smokers (quit <2y ago) are defined differently than the main sample CA, EN, US former smokers (quit between 2-5y ago).*

** Core sample (a subsample of the main sample) (see Section 6) = Records for respondents from CA, EN, and US who quit >2y ago have been removed from the main sample so that the former smoker samples in each country meet the same definition. *The core sample should be used for analyzing former smoker samples across AU and other countries.*

3.2 Sample definition and eligibility (inclusion/exclusion) criteria

- Inclusion/exclusion criteria were the same in Canada, England, and the US. In Australia, however, the criteria were different (see Table 4.1, Table 4.2).
- The sample in each country is specified on three dimensions: Source, cohort status, and user type (by age). These statuses determine fieldwork procedures for each subsample.

Sample Dimension 1: Source

1) ITC-owned cohort:
Sample contact information has been provided directly to ITC. During 4CV5, SCSRU and/or ITC will contact ITC-owned sample directly.

2) Panel-owned:
Sample is provided by a panel firm and may include one or both of cohort and replenishment subsamples. At 4CV5, panel firms will contact panel-owned sample. ITC/SCSRU are not permitted to contact the sample directly. shou

- **Firms** varied in their **quality of samples**. Some firms had high-quality probability-based samples, and some had opt-in samples, and others a combination of these. For this reason, **approximate quota targets (**based on region, sex, and/or age) were established proportional to national benchmarks to ensure appropriate distributions were realized in the final sample.

Sample Dimension 2: Cohort status

1) Recontact (cohort) sample:
All previous 4CV4 respondents are recontact sample. Inclusion criteria (Table 4.1, Table 4.2) and procedures differ for the cohort sample compared to the replenishment sample.

2) Replenishment (fresh) sample:
New recruits at 4CV5 make up the replenishment/fresh sample. This sample is recruited to replace respondents (from the recontact sample) who were lost to follow-up. Inclusion criteria (Table 4.1, Table 4.2) and procedures differ for the cohort sample compared to the replenishment sample.

Sample Dimension 3: User type

1) Current at-least-monthly-cigarette smoker or a **recent former smoker** (i.e., smoked at least monthly previously, quit in the past 24 months), aged 18y or older

2) Current at-least-weekly NVP/HTP/snus/Tobacco-free Oral Nicotine Product (TF-ONP) user, aged 18y or older

Notes:

- Eligibility inclusion/exclusion criteria for the survey is based on user type (criteria for Australia was different vs. Canada, England, US) and age, by cohort status vs. replenishment status (Table 4.1, Table 4.2).
- In AU, NVP users only are sampled in addition to smokers/recent former smokers.
- In CA, EN, US, **snus and TF-ONP users** are newly eligible at Wave 5, in addition to NVP/HTP users.
- **Multi-product users** (smokers who also use NVPs/HTPs/snus/ONPs at least weekly) will be recruited as part of the current NVP/HTP/snus/TF-ONP user quotas or, when full, as part of the cigarette smokers/recent former smokers quotas.

- **Inclusion/exclusion criteria** differed between the **Recontact** and the **Replenishment** samples (Table 4.1, Table 4.2).

Table 4.1. Inclusion/exclusion criteria for 4CV5 Canada, England, and US *(but not Australia)*.

Cohort Sample – CA, EN, US only

All past ITC-owned and panel-owned 4CV4 respondents were eligible, irrespective of current status, **except if:**

1) they were recruited at or prior to 4CV1 and had not smoked at least monthly or used ANDs as specified below in the last 5 years,
2) they provided poor quality data at 4CV4 or someone else was known to have completed the survey, or
3) they failed the age check.

Cohort sample eligibility (CA, EN, US only)

Recruitment Wave status	Screening at 4CV5	4CV5 eligibility
Recontact or Replenishment sample at 4CV1 (Jul–Oct 2016); 4CV1 status was short-term (quit<24 months ago) or long-term quitter (Recontact only, quit >24 months ago) and non-vaper (vapes not-at-all, less-than-monthly, or monthly).	(Smoked at least monthly at any time between and including 4CV1 to 4CV5), OR (Vaped at least weekly at any time between and including 4CV1 to 4CV5), OR (Used HTP at least weekly at any time between and including 4CV3 to 4CV5), OR (Used snus or TF-ONPs at least weekly at any time between 4CV4 to 4CV5)	Eligible
Recontact or Replenishment sample at 4CV1 (Jul–Oct 2016); 4CV1 status was short-term (quit<24 months ago) or long-term quitter (quit >24 months ago) and non-vaper (vapes not-at-all, less-than-monthly, or monthly).	Opposite of above cell. That is… (Smoked less than monthly or not at all at all between 4CV1 to 4CV5), AND (Vaped at less than weekly or not at all at all between and including 4CV1 to 4CV5), AND (Used HTP less than weekly or not at all at all between and including 4CV3 to 4CV5), AND (Used snus or TF-ONPs less than weekly or not at all at all between and including 4CV4 to 4CV5)	Ineligible
Recontact sample at 4CV2 (Jul-Oct 2018); 4CV2 status was recent cigarette quitter (quit <24 months ago).	Irrespective of 4CV5 status	Eligible
Recruited as replenishment sample at 4CV2 or 4CV3	Irrespective of 4CV5 status	Eligible

Note: Smokers who also use NVPs/HTPs/snus/TF-ONPs at least weekly were recruited either as part of the current NVPs/HTPs/snus/TF-ONPs user sample or as part of the smoker/ quitter sample. Programming counted these cases toward the NVPs/HTPs/snus/TF-ONPs sample first and, when full, toward the smoker/quitter sample second.

Replenishment Sample – CA, EN, US only

Met age, sex, and/or region criteria and user type criteria (below) for an open quota target.
User type criteria were:

1) **Current smoker.** Smoke at least monthly, had smoked at least 100 cigarettes; irrespective of NVP/HTP/snus/TF-ONP use.
2) **Recent former smoker.** Quit smoking in the last 24 months, had smoked at least 100 cigarettes in lifetime; irrespective of NVP/HTP/snus/TF-ONP use.
3) **Current NVP, HTP, snus, or tobacco-free oral nicotine product (TF-ONP) user.** Use NVPs/HTPs/snus/TF-ONPs at least weekly; irrespective of cigarette smoking.

Replenishment sample eligibility (CA, EN, US only)

Subsample group	Definition	Definition using 4CV4 variables
Smoker OR recent quitter	[Currently smokes cigarettes at least monthly AND smoked at least 100 cigs in lifetime] OR [currently smokes not at all, has quit in the past 24 months, and has smoked at least 100 cigarettes in lifetime]	FR309v=10-31 OR FR309v=40
At-least-weekly NVP, HTP, snus, TF-ONP user	[Currently uses NVP at least weekly] OR [currently uses HTP devices at least weekly] OR [currently uses snus at least weekly] OR [currently uses TF-ONPs at least weekly]	EC309v=10-20 OR HN309v=10-20 OR NC006=1-2 OR BI038=1-2

Note: Current at-least-monthly smokers/quitters in the last 24 months who also use NVPs/HTPs at least weekly (i.e., meet definitions for both Smoker/Recent quitter and NVP/HTP user) will be recruited from the same country-specific firms. For new recruits who meet criteria for both subsample groups, they will initially be attributed to the NVP/HTP user sample targets or, when full, will be attributed to the Smoker/Recent quitter sample targets.

Note: Each respondent had a status for: smoking, vaping, HTP use, snus use, and TF-ONP use. Thus, many combinations of the above user types, as well as non-users of one product who use -- or had used -- the other product(s) were possible in the sample.

Table 4.2. Inclusion/exclusion criteria for 4CV5 *Australia only*.

Cohort Sample – Australia only

All past ITC-owned and panel-owned 4CV4 respondents were eligible, irrespective of current status, **except if:**

1) they were recruited at or prior to 4CV1 and had not smoked at least monthly or used NVP/HTP at least weekly in the last 5 years,
2) they provided poor quality data at 4CV4 or someone else was known to have completed the survey, or
3) they failed the age check.

Cohort sample eligibility (Australia only)

Recruitment Wave status	Screening at 4CV5	4CV4 eligibility
Recontact or Replenishment sample at 4CV1 (Jul–Oct 2016); 4CV1 status was short-term (quit<24 months ago) or long-term quitter (Recontact only, quit >24 months ago) and non-vaper (vapes not-at-all, less-than-monthly, or monthly).	(Smoked at least monthly at any time between and including 4CV1 to 4CV5), OR (Vaped at least weekly at any time between and including 4CV1 to 4CV5), OR (Used HTP at least weekly at 4CV3),	Eligible
Recontact or Replenishment sample at 4CV1 (Jul–Oct 2016); 4CV1 status was short-term (quit<24 months ago) or long-term quitter (quit >24 months ago) and non-vaper (vapes not-at-all, less-than-monthly, or monthly).	Opposite of above cell. That is… (Smoked less than monthly or not at all at all between 4CV1 to 4CV5), AND (Vaped at less than weekly or not at all at all between and including 4CV1 to 4CV5), AND (Used HTP less than weekly or not at all at all at 4CV3).	Ineligible
Recontact sample at 4CV2 (Jul-Oct 2018); 4CV2 status was recent cigarette quitter (quit <24 months ago).	Regardless of 4CV5 status	Eligible
Recruited as a replenishment sample at 4CV2 through 4CV4	Regardless of 4CV5 status	Eligible

Replenishment Sample – Australia only

Met age, sex, and user type criteria (below) for an open quota target.

User type criteria were:

1) **Current smoker.** Smoke at least monthly, had smoked at least 100 cigarettes lifetime; irrespective of NVP use.
2) **Recent former smoker.** Quit smoking in the last 24 months, had smoked at least 100 cigarettes in lifetime; irrespective of NVP use.
3) **Current NVP user.** Use NVPs at least weekly; irrespective of cigarette smoking.

Replenishment sample eligibility (Australia only)

Subsample group	Definition	Definition using 4CV4 variables
Smoker OR recent quitter	[Currently smokes cigarettes at least monthly AND smoked at least 100 cigarettes in lifetime] OR [currently smokes not at all, has quit in the past 24 months, and has smoked at least 100 cigarettes in lifetime]	FR309v=10-31 OR FR309v=40
At-least-weekly NVP user	Currently vaping at least weekly	EC309v=10-20

Note: Each respondent had a status for: smoking, vaping, HTP use, snus use, and TF-ONP use. Thus, many combinations of the above user types, as well as non-users of one product who use -- or had used -- the other product(s) were possible in the sample.

ETM Sample

Main survey Recontact and Replenishment panel-owned sample from CA, EN, or US who were aged 18y or older, and met criteria for one of three user categories:

1) **Smoker only.** Daily cigarette smoker, smoked factory-made cigarettes more than or equal to roll-your-own, not NVP user.
2) **Daily smoker/ weekly vaper.** Daily cigarette smoker, smoke factory-made cigarettes more than or equal to roll-your-own, weekly NVP users.

4 SURVEY QUESTIONNAIRE

The survey questionnaire was developed by ITC investigators in collaboration with the project management and survey management teams. Although a core set of questions is included in all 4CV waves; at each wave, the questionnaire is updated to best address the research questions in changing market, environment, and policy contexts while also improving survey respondents' experience as much as possible.

4.1 4CV5 Survey Development Process and Timeline

The ITC survey development process comprises two main phases: 1) Review of survey questions, and 2) Operationalization, with some overlap between phases.

Table 5. 4CV5 Survey development timeline

Survey review	Survey operationalization
Round 1: Mar 20, 2023 – July 15, 2023 **Round 2**: Sep 28, 2023 – Dec 21, 2023 **Round 3**: Jan 31, 2024 – Apr 30, 2024	May 15, 2023 – Jul 15, 2024

Review of survey questions

During Round 1, five working groups were established for survey content areas based on team members areas of expertise and interest. Each group systematically **reviewed the previous wave's questionnaire** content and made a formal proposal to the larger expert group to delete, add, or revise content. Large-group discussions and review were held during Round 2 where additions and revisions were considered to the survey. Round 3 consisted of deletions to ensure that the survey met the required length of interview.

Survey operationalization

After Round 2 of the systematic review of questions within Phase 1 was completed, the survey draft was sent to the ITC Survey Management Group (SMG) for **operationalization** (Phase 2). Phase 2 involved comprehensively and iteratively reviewing and revising the survey to ensure that routing (i.e., use of filters and skips to show proper set of questions to different users), question wording, response options, and all other survey elements were refined and cross-referenced for consistency, clarity, and accuracy for programming. At the conclusion of Phase 2, the final draft of the survey was generated by SMG and sent to the survey firm for programming and testing.

4.2 4CV5 Survey Content

To update the survey at 4CV5, the previous 4CV4 Survey content was reviewed and revised to delete items no longer of interest or useful in analyses, add items measuring constructs of emerging interest, and revise existing items for improved measurement as applicable (Table 6). Survey content was rearranged at 4CV5 to accommodate newer content and yet still maintaining comparability to previous 4CV surveys.

- Of the survey content in 4CV5, about 71.6% was from the 4CV4 Survey. About 23.6% new content was added at 4CV4.
- Some new content was asked of only a small proportion of the sample.
- At 4CV5, cohort participants were no longer separated into module (Regmod1 or regmod2) for Cigarette Regulation and for Vaping Regulation as they had been in Waves 3 and 4.
- Additions at 4CV5 included:
    - Defining (and screening for eligibility on at-least-weekly use of) snus and tobacco-free oral nicotine products; and asking about reasons for use, risk perceptions, and hypothetical scenarios under which respondents would use these products.

    - Updated questions on cannabis use and co-use with nicotine products.
- The median survey length was **44 minutes**.
- A copy of the survey instrument is available on request or at http://itcproject.org/surveys.

Table 6. Measures in the 4CV5 Survey

Demographic Variables: Sex, gender, age, ethnicity, education, income, state of health
Other personal moderators: Quitting history, nicotine dependence, levels of stress including financial stress and depressed mood, use of intoxicants (*e.g.,* alcohol, cannabis), time perspective, etc.
Environmental moderators: Number of smokers/users in household, and in social network
Policy-specific (proximal) measures of FCTC policies (cigarettes) and policies on NVPs/HTPs: 1) **Article 6 (cigarettes, NVPs, HTPs):** Price paid per unit of product, total weekly cost, product type/variant, purchasing unit, price perceptions. 2) **Article 14 (cigarettes, NVPs, HTPs):** Use of cessation services & recall of advice, use of NVPs/HTPs/other ANDs and/or other medicines in conjunction with professional assistance, advice on appropriateness of NVP/HTP use. 3) **Article 13 (cigarettes, NVPs, HTPs):** Advertising/ marketing: noticing ads and frequency in key channels (TV, print, internet), susceptibility to advertising, whether NVP/HTP advertising makes respondents think about cigarettes 4) **Article 11 (cigarettes, NVPs, HTPs):** Health warnings: salience and noticing of health warnings (if any), brand usage, perceived risks, perceived impact on product use; forgoing cigarettes because of warnings. 5) **Article 8 (cigarettes, NVPs, HTPs):** Smoke-free/vapour-free laws (and/or established policies), exposure to smoking/vaping in key venues, perceived impact laws/policies on product use, reports on restrictions. 6) **Product availability (NVPs/HTPs/snus/TF-ONPs)**: Restrictions on access, perceived availability. 7) **Article 9 (mostly NVPs/HTPs):** Nicotine content, flavour and other product characteristics: nicotine content and flavours of NVP/HTP brands used, perceived addictiveness of cigarettes/NVPs/HTPs, and appeal of NVPs/HTPs. 8) **Article 12 (all)**: Awareness/recall of media campaigns or information on NVPs/HTPs, and on anti-smoking themes.
Psychosocial mediator (distal) variables: Knowledge/awareness/beliefs about harmfulness of cigarettes/NVPs/HTPs/snus/TF-ONPs and nicotine overall, social norms for cigarettes/NVPs/HTPs, outcome expectancies for products, reasons for using NVPs/HTPs/snus/TF-ONPs, self-efficacy and intentions to quit smoking; perceived harmfulness of NVPs/ HTPs/snus/TF-ONPs relative to cigarettes, health concerns, functions of smoking and nicotine use.
Tobacco/nicotine use behaviors **(Cigarettes, NVPs, HTPs):** History and current levels of use: frequency, duration, intensity of use (*e.g.,* cigarettes per day), usual brand/type of product; history of smoking quit attempts, use of aids, targeted questions about last quit attempt (timing, length, aids used, duration of abstinence, reasons for success/relapse); product switching and reasons for switching (and reasons for starting/quitting NVPs/HTPs). **(Emerging ANDs, including snus and TF-ONPs):** Current frequency of use, brands used.

5 Survey Firms

Table 7 summarizes the survey firms and additional teams involved in the 4CV5 fieldwork.

Table 7. 4CV5 Stakeholders and Responsibilities

Country	Project Management	Programming & Web Hosting	Recontact & Replenishment	Processing Remuneration	Data Management
CA	ITC PMG	SCSRU	SCSRU, Leger	ITC Admin Group (ITC-owned) Leger (Panel-owned)	ITC DMC
EN	ITC PMG	SCSRU	SCSRU, Ipsos UK*, Rakuten**	Logistic Solutions (ITC-owned) Ipsos UK*, Rakuten** (Panel-owned)	ITC DMC
US	ITC PMG	SCSRU	SCSRU, Ipsos PA, Ipsos UK	ITC Admin Group (ITC-owned) Ipsos UK* (panel-owned), Ipsos PA (panel-owned)	ITC DMC
AU	ITC PMG	SCSRU	RM**	RM** (ITC-owned) RM** (Panel-owned)	ITC DMC

* Previously known as Ipsos MORI; ** With panel partner(s)

5.1 Description of country survey firms' recruitment procedures

Firms varied in their **quality of sample**. Some firms had high-quality probability-based samples, and some had opt-in samples, and others a combination of these. For this reason, **approximate quota targets (**based on region, sex, and/or age) were established proportional to national benchmarks to ensure appropriate distributions were realized in the final sample.

Each survey firm provided below a description of their panel as it was at the time of the 4CV5 Survey.

5.1.1 Australia - Roy Morgan (RM) and Dynata (previously Survey Sampling International, SSI)

Sample: Australian smokers, recent former smokers, and NVP users.

RM provided the following description for their relevant panels at the time of the 4CV5 Survey: **RM**'s tasks in this project were carried out in compliance with ISO 20252 Market, Opinion & Social Research. RM's proprietary survey **Single Source** is representative of the Australian population aged 14y+ in terms of sex, age and geographical location. Sample selection was conducted for Single Source via telephone interviewing using a randomized cluster sampling approach for household selection and a rule of priority approach for respondent selection within the household. Rigorous sampling procedures were applied each month to ensure that respondents reflect the key demographic characteristics of the Australian population. Interviewing for Single Source is conducted weekly, so our sample was replenished continuously.

This survey method also serves as a form of validation for the respondent – we know that someone who claims to be "male, aged 29 years, living in Tamworth, NSW" is indeed that, because we have spoken with him and interviewed him. For 4CV5, RM drew Single Source respondents who were over 18y, interviewed since December 2022 (when fieldwork was conducted for ITC 4CV4), had provided a telephone number for recontact and indicated during their Single Source interview that they smoked factory-made cigarettes (FMC) or NVPs at that time. Those who recently quit smoking would be identified during the computer-assisted telephone interview (CATI) screening survey. By recontacting Single Source respondents who broadly qualified for 4CV, we were able to ensure that the sample would be representative of the Australian smoking population. The sample was then screened via CATI to ensure it met UW's specifications and were automatically sent an email invitation from the telephone interviewing system, once they qualified for the research and had provided an email address.

Dynata (previously SSI) was RM's partner to obtain the 4CV5 sample in AU. Dynata panelists were recruited via partnerships with other corporations or invited by banners, invitations and messages. To minimize the impact of different partnerships on survey results, Dynata uses a combination of personality and psychographic characteristics to understand and identify the underlying traits which make a difference in the way people answer survey questions. By asking participants a short set of key questions, Dynata can control the characteristics of people within the sample and allows Dynata to provide an exceptionally consistent sample blend. Dynata also uses digital fingerprinting, to ensure that the same person does not take a survey more than once from the same device.

The CATI recruitment survey for Single Source sample was converted to an online format for Dynata panelists and collected additional personal information (email address, physical address, and best contact number). The questions were phrased in exactly the same manner, although the response options were slightly modified to better suit an online environment. When a Dynata panelist completed the online recruitment survey, their details were transferred into the Roy Morgan redirect survey. Once an hour the redirect survey automatically sent out an invitation to the 4CV4 Survey using the email format provided by UW.

Snowball Sampling was utilized to fill a shortfall of 18-24year-old respondents (approximately 32 at the time of implementation) toward the end of fieldwork. This age group is a valuable part of the sample. Roy Morgan (and Dynata) had exhausted all potential samples in the 18-24-year-olds. Upon discussion with the University of Waterloo and University of Queensland, it was agreed to obtain the final 32 responses via a snowball sample methodology. Past respondents were approached, in a tiered approach starting with 18-24-year-olds and then 25+ after if needed, to nominate up to 5 respondents who were 18-24 years old and either smoke or vape (or both). Nominated respondents were sent a text message enquiring whether they would participate in the survey, along with a link to a screening survey to ensure eligibility for the 4CV5 AU survey.

5.1.2 Canada - Leger

Sample: Canada smokers, recent former smokers, and NVP/HTP/snus/TF-ONP users
Respondents were selected at random from the Leger web panel to participate in the study.

Leger provided the following description for their panel at the time of the 4CV5 Survey:
Leger's online panel has more than 400,000 members nationally and has a retention rate of 90%. The Leger panel is high-quality because members have largely been sourced from samples considered nationally representative of Canada.

5.1.3 England

5.1.3.1 *Rakuten Insight*

Sample: England smokers, recent former smokers, and NVP/HTP/snus/TF-ONP users
Rakuten Insight's sample in England was recruited through one main panel partner, Dynata (previously known as SSI), Pureprofile (via Dynata), Market Cube (via Dynata), and Panelbase.

Rakuten Insight provided the following description for their relevant panels at the time of the 4CV5 Survey: Rakuten Insight works very closely with its panel partner in UK, which has a variety of sample sources such as panels, web intercept samples, and specialty lists. Each of the recruitment channels delivers a different population and slightly different results, thus increasing diversity and representativeness. The loyalty panels are actively managed, high-quality online access panels built from two decades of experience. Invitation-only recruitment campaigns are run via direct email and achieve this by partnering with over 75 globally recognized consumer and business-facing brands. Additionally, the panels are localized – not just translated – with native language panel support and country-specific reward choices. This ensures participants are attentive and content, resulting in accurate, thoughtful answers and, ultimately, high-quality data. In view of the scope and scale of the 4CV research, two additional approved panel partners, which met Rakuten Insight's quality control procedures, were also used in England to support with the project.

5.1.3.2 *England - Ipsos UK (previoulsy Ipsos MORI)**

Sample: England smokers, recent former smokers, and and NVP users
**Ipsos UK also recruited a sample of US smokers and recent former smokers and NVP users. The same description applies for both samples.*

Ipsos UK provided the following description for their relevant panels at the time of the 4CV5 Survey: Ipsos UK's online panels are subject to rigorous recruitment procedures aimed at ensuring accuracy, consistency and non-duplication. To join, panel applicants are validated by a means of sophisticated vetting procedures, using a variety of opt-in recruitment channels. Shortly after joining, panel members' survey-taking behavior is tested, with those most likely to make intentional or unintentional errors on future surveys deactivated. Subsequently, panelists' behavior is monitored and tracked across all surveys. Ipsos UK employs purging procedures based on these data to remove suspects from eligible sampling pools. In view of the scope and scale of the 4CV research, approved panel partners, that met Ipsos UK quality control procedures, were also used in EN (but not in the US) to support with the project.

5.1.4 United States

5.1.4.1 *Ipsos KnowledgePanel® - Authorized Language for Client Involving Descriptions of KnowledgePanel® Methodology*

Sample: US smokers, recent former smokers, and and NVP/HTP/snus/TF-ONP users
The survey was conducted using the web-enabled KnowledgePanel®, a probability-based panel designed to be representative of the U.S. population. Initially, participants are chosen scientifically by a random selection of telephone numbers and residential addresses. Persons in selected households are then invited by telephone or by mail to participate in the web-enabled KnowledgePanel. For those who agree to participate, but do not already have Internet access, Ipsos provides at no cost a tablet and data plan. People who already have computers and Internet service are permitted to participate using their own equipment. Panelists then receive unique log-in information for accessing surveys online, and then are sent emails throughout each month inviting them to participate in research.

Ipsos Public Affairs

Ipsos PA provided the following description for their relevant panels at the time of the 4CV5 Survey: Ipsos is an independent market research company controlled and managed by research professionals. Founded in France in 1975, Ipsos has grown into a worldwide research group with a strong presence in all key markets. Ipsos ranks third in the global research industry. At Ipsos we are passionately curious about people, markets, brands and society. We make our changing world easier and faster to navigate and inspire clients to make smarter decisions. We deliver with security, speed, simplicity and substance. We are Game Changers. With offices in 88 countries, Ipsos delivers insightful expertise across six research specializations: advertising, customer loyalty, marketing, media, public affairs research, and survey management. Ipsos researchers assess market potential and interpret market trends. We develop and build brands. We help clients build long-term relationships with their customers. We test advertising and study audience responses to various media and they measure public opinion around the globe. Visit www.ipsos.com/en-us to learn more about Ipsos' offerings and capabilities.

5.1.4.2 *Ipsos UK (previously Ipsos MORI)**

Sample: US smokers and recent former smokers and NVP users

***Ipsos UK also recruited a sample of England smokers, former smokers, and NVP users. The same description applies for both samples** (see Section 5.1.3.2 for the panel description provided by **Ipsos UK**).

5.1.5 Experimental Tobacco Marketplace (ETM)

5.1.5.1 *Virginia Tech*

As a sub-study of 4CV5 run out of Virginia Tech (VT), they were responsible for recruiting ETM respondents from Canada, England, and the US. VT obtained ethics clearance for the sub-study, as well as programmed and hosted the ETM. This included setting up URLs to redirect from the survey to the ETM sub-study.

6 Study Procedures

6.1 Sample criteria and definitions used in fieldwork

- The sample in each country was specified on three dimensions: Source, cohort status, and user type (described in Section 3.2). The statuses determined fieldwork procedures for each subsample.

6.2 Phase 1 (Recontact) and Phase 2 (Replenishment) of surveying

- Phase 1, Recontact: The initial phase of fieldwork includes inviting Recontact sample to complete the 4CV5 Survey. After an appropriate time period, a valid retention estimate will be determined. Once retention has plateaued, the Phase 2, Replenishment fieldwork will start.
- Phase 2, Replenishment: Begins when fresh sample are invited to the survey to replace those lost to follow up. Replenishment fieldwork often overlaps with Recontact fieldwork after the initial lead in period.

6.3 Pre-fieldwork procedures for all countries

6.3.1 Ethics review and service agreements

- Study materials and procedures were reviewed and cleared by Research Ethics Committees at all study institutions, with the exception of those that waived ethics review, as described in Section 2.3, prior to initiating fieldwork.
- Service Agreements were executed between all service providers and the University of Waterloo (study lead institution for CA, EN, and US) or the University of Queensland (lead institution for AU) prior to initiating fieldwork.

6.3.2 Programming and testing the 4CV5 Survey and ETM

- The 4CV5 Survey content was provided to firms for review and clearance prior to initiating fieldwork.
- SCSRU programmed and tested the 4CV5 Survey extensively to ensure the system functioned as intended, in collaboration with ITC and panel firms prior to fieldwork.
- VT programmed and tested the ETM to ensure the system functioned as intended, in collaboration with ITC, and panel firms.

6.3.3 Determining Recontact Sample

- The ITC DMC provided data files listing ITC uniqIDs and PanelIDs of valid Recontact sample (i.e., 'speeders' omitted, unsubscribes omitted) to each appropriate survey firm. These people were recontacted and invited to 4CV5.
- Relevant 'smart data' fields for each valid cohort ID were read-into the 4CV5 Survey for routing purposes and/or to determine eligibility at 4CV5.

6.3.4 Predicting Replenishment targets

- ITC worked with panel firms to predict retention and determine initial Replenishment sample estimates prior to the beginning of fieldwork.
- Replenishment target estimates were then monitored and updated during fieldwork as necessary, depending on the actual number Recontact surveys completed (Appendix B).

6.3.5 Technical set up to direct respondents to and from the survey

- Unique personal identification numbers (PINs) identified survey records in the online survey system. PINs were linked to panel-specific ID numbers and ITC-specific ID numbers in secure files, in order to track participation and remunerate participants, while maintaining confidential survey response data.

    - For Recontact sample, PINs were pre-assigned by SCSRU and provided to relevant panel firms.
    - For Replenishment sample, PINs were assigned by the panel firm at the time of recruitment (linkage information was transferred to ITC after fieldwork was completed).
- URLs were used to direct panelists from their home firms to the 4CV5 Survey (hosted at UW) and the ETM (hosted by VT) if relevant.
- Redirecting URLs directed panelists back to their home panel's website in order to be processed as 'survey/ETM completes' and remunerated, or to be processed as 'ineligible' for the survey/ETM.

6.4 Fieldwork procedures - AU ITC-owned and RM-owned Recontact sample

- ITC-owned and RM-owned sample from AU were managed by RM; these respondents met the following definitions:
    - **ITC-owned Recontact sample** had originally been recruited by ITC as current smokers from an RDD sampling frame, had provided their contact information, and agreed to receive requests for participation in future surveys from ITC.
    - **RM-owned Recontact sample** had originally been recruited through RM's Single Source panel since 4C9 or through a panel partner.
- RM managed all communications with their respective panelists, and directed eligible and willing panelists to the 4CV5 Survey.
- ITC provided to RM a list of all eligible AU sample to be recontacted and invited to participate in the follow-up 4CV5 Survey. Past respondents who had provided poor-quality data at 4CV4, who had requested no future contact, or who were no longer on panel, were not recontacted. Any other eligibility criteria were applied during screening at the beginning of the survey.

6.4.1 AU ITC-owned and panel-owned Recontact sample: Interwave reminders, survey invitations, and survey reminders

- ***Interwave reminders*** were sent to eligible sample before fieldwork started. The interwave reminders highlighted the importance of the survey and described the time commitment and remuneration value (Table 8). RM's invitations used the ITC logo, a picture of the investigator team, and ITC phrasing that described the importance of the survey, the time commitment and remuneration value, and provided a personalized direct link to the survey.
- ***Email survey invitations*** were sent to eligible Recontact sample, in batches, beginning on Day 1 of fieldwork. RM's invitations used the ITC logo, a picture of the investigator team, and ITC phrasing that described the importance of the survey, the time commitment and remuneration value, and provided a personalized direct link to the survey.
- ***Email reminders*** were sent to all panelists who had been invited but not yet completed the survey on Days 3, 8, 10, and 15, as needed (note: Day 1 is the day the initial email incitation was sent). An additional email reminder was sent a few days prior to fieldwork closing to alert any panelists who wished to submit their surveys.
- RM's contact information was provided in email invitations and reminders, and in the survey, so that cohort panelists could contact the panel or to opt out. Contact information for the principal investigator and an ethics contact person was provided in the survey, and in some email invitations/reminders.

6.4.2 AU ITC-owned and panel-owned Recontact sample: Phone reminders

- In order to improve the survey response, phone reminders were conducted using a pre-determined calling priority list:
    - Those who had provided an email address, but the 4CV5 email invitation bounced back.
    - Those who had not accessed the web survey after 1 week from the date of the email reminder.

- A maximum of 7 call attempts were made; on the 7th call attempt, the interviewer left an appropriate voicemail message, and then no further attempts were made to contact the intended participant.
- Standard best practices were followed to ensure safety and confidentiality.

6.4.3 AU ITC-owned and panel-owned Recontact sample: Remuneration

- Remuneration was dependent on completing the 4CV5 Survey. See Table 8 for values.

6.4.4 AU ITC-owned and panel-owned Recontact sample: Technical support

- RM's contact information was provided in recruitment emails and the 4CV5 Survey, for technical support and/or comments.
- RM's toll-free number for was also provided in recruitment emails and the 4CV5 Survey, and a staff person was designated to check messages on the toll-free number on a daily basis during fieldwork.
- RM responded to any queries within 48 hours.

6.5 AU Replenishment Procedures—Panel-owned Respondents

- To account for attrition and maintain the target sample size, new respondents were recruited at 4CV4.
- ITC DMC provided RM with updated replenishment targets, at appropriate intervals during the fieldwork period as the actual retention rate (of Recontact sample) became known.
- The replenishment sample was recruited from RM Single Source (via CATI), Dynata (online), and via snowball recruitment (18-24-year-olds) prior to being redirected to the 4CV5 Survey.
- RM/Dynata emailed their panelists who met preliminary eligibility criteria and provided them with a personalized link to the 4CV4 Survey.
- RM was in charge of the snowball recruitment procedure.
- The final screening and allocation to quotas took place within the survey.
- **Preliminary eligibility criteria** were:
  - Any panelist who had reported being a smoker or former smoker at any panel update in the past 24 months.
  - Any panelist who had reported being a NVP user who screens as vaping at least weekly.

6.5.1 AU Replenishment sample quota targets

- Appendix B.1 describes the AU Replenishment sample quota targets.

6.5.2 AU panel-owned Replenishment sample: Interwave reminders, survey invitations, and survey reminders

- ***Email survey invitations*** were sent to panel-owned Replenishment sample who met preliminary criteria in batches until the Replenishment targets with the corresponding sample characteristics were attained. RM's invitations used the ITC logo, a picture of the investigator team, and ITC phrasing that described the importance of the survey, the time commitment and remuneration value, and provided a personalized direct link to the survey.
- ***Email reminders*** were sent to all panelists who had been invited but not yet completed the survey on Days 3, 8, 10, and 15, as needed (note: Day 1 is the day the initial email incitation was sent). An additional email reminder was sent a few days prior to fieldwork closing to alert any panelists who wished to submit their surveys. Email reminders used the ITC logo, a picture of the investigator team, and ITC phrasing that described the importance of the survey, the time commitment and remuneration value, and provided a personalized direct link to the survey.
- RM's contact information was provided in email invitations and reminders, and in the survey, so that cohort panelists could contact the panel or to opt out. Contact information for the principal

investigator and an ethics contact person was provided in the survey, and in some email invitations/reminders.

6.5.3 AU panel-owned Replenishment sample: Remuneration procedures

- Remuneration was dependent on completing the 4CV5 Survey. See Table 8 for values.
- Additional remuneration was provided to past respondents who successfully nominated an 18-24 year old participant during the snowball recruitment period.

6.5.4 AU panel-owned Replenishment sample: Technical support

- RM's contact information was provided in recruitment emails and the 4CV5 Survey, for technical support and/or comments.
- RM's toll-free number for was also provided in recruitment emails and the 4CV5 Survey, and a staff person was designated to check messages on the toll-free number on a daily basis during fieldwork.
- RM responded to any queries within 48 hours.

6.6 Fieldwork Procedures - CA, EN, and US ITC-owned Recontact sample

- All ITC-owned sample are Recontact sample
- ITC-owned sample from CA, EN, and the US were managed by the SCSRU; these respondents met the following definitions:
  - ***ITC-owned 4C cohort***: ITC-owned respondents who participated at 4CV4 and who were originally recruited *from a telephone RDD sampling frame* at their first 4C Survey, and
  - ***New ITC-owned***: Respondents recruited from a US opt-in panel partner at 4CV1 or 4CV2.
- ITC provided to SCSRU a list of all eligible ITC-owned sample to be recontacted and invited to participate in the follow-up 4CV5 Survey.
  - Past respondents who had provided poor-quality data at 4CV4, who had requested no future contact, or whose contact information were outdated, were not recontacted. Any other eligibility criteria were applied during screening at the beginning of the survey.

6.6.1 CA, EN, and US ITC-owned Recontact Sample: Interwave reminders, survey invitations, and survey reminders (email vs. letter)

- As of the completion of Wave 4, all communications with the cohort are now via email, as following Wave 4, no respondent remains in the sample who has not provided an email address.
  - Cohort members were sent a pre-fieldwork interwave email reminder, an email invitation, and were placed on an email reminder schedule. Some of them also qualified for a reminder call.
- Contact information for the SCSRU, an ethics committee representative, and a study investigator were provided in all invitations, reminders, and in the survey. Cohort members could phone or send a message to update their contact information or to opt out of the study.
- ***Interwave reminders*** were sent eligible sample approximately 2 weeks before the fieldwork start date. The reminders highlighted the importance of the survey and reminded them of the remuneration that they would get for submitting a complete survey.
- ***Email survey invitations*** were sent to eligible ITC-owned cohort members (with an email on file), in batches, beginning on Day 1 of fieldwork. Invitations described the study and invited the respondent to participate by clicking on a personalized direct link to the survey.
- ***Email reminders*** were sent to all ITC-owned cohort members who had been invited but not yet completed the survey on Days 3, 8, 10, and 15, (*Note: Day 1 is the day the initial email invitation is sent).*
- About one week before the survey was to close, one last reminder email was sent to all outstanding invited respondents who had not submitted a completed survey **IF** they had not responded stating that they did not wish to participate/wished to withdraw.

6.6.2 CA, EN, and US ITC-owned Recontact sample: Phone reminders

- **3-4 weeks after the start of fieldwork**, SCSRU implemented phone reminders based on a pre-established calling priority list:
  - Those who did not provide an email address at 4CV4 and completed last wave by phone.
  - Those who provided an email address, but the 4CV5 email invitation bounced back, and completed last wave by phone.
  - Those who provided a valid email address and completed last wave by phone.
  - Those who provided an email address, but the 4CV5 email invitation bounced back, and completed last wave by web, including those from the new ITC-owned group.
  - Those who received a reminder call and completed the survey by web later during fieldwork at 4CV4.
  - Those who had not accessed the web survey after 3-4 weeks in the field, including those from the new ITC-owned group.
  - Partial completes.
- A maximum of 7 call attempts were made. On the 1st and 7th call attempts, the interviewer left an appropriate voicemail message.
- SCSRU provided guidelines to their interviewers on how to call respondents on cell phones to ensure safety and confidentiality.

6.6.3 CA, EN, and US ITC-owned Recontact sample: Remuneration

- Remuneration was dependent on completing the 4CV5 Survey. See Table 8 for values.

6.6.4 CA, EN, US ITC-owned Recontact sample: Technical support

- SCSRU's contact information appeared both in the correspondence to the respondents and in the survey.
- Respondents from each country were able to send email messages to the SCSRU for technical support requests and/or comments.
- There was only one toll-free number servicing North America (and no toll-free number for England) at Wave 5. SCSRU designated a staff person to check messages on the toll-free number on a daily basis during the fieldwork period and respond to messages as needed.

6.7 Fieldwork Procedures – CA, EN, and US Panel-owned Recontact sample

- The **panel-owned Recontact sample** included panelists recruited into the study by one of the following survey firms at 4C9-10 or 4CV1-4:
  - CA: Leger
  - US: Ipsos PA*; Ipsos UK (previously Ipsos MORI)
  - EN: Ipsos UK (previously Ipsos MORI), Rakuten

  **Respondents recruited from Ipsos PA's KnowledgePanel are "panel-owned". Those recruited through the US opt-in panel partner at 4CV1 and 4CV2 were transferred to ITC, and hence were "ITC-owned" at subsequent waves.*
- Panel firms managed all communications with their respective panelists and directed eligible willing panelists to the 4CV5 Survey and ETM.
- ITC provided to each panel firm a list of all eligible panel-owned sample to be recontacted and invited to participate in the follow-up 4CV5 Survey.
  - Past panelist respondents who had provided poor-quality data at the previous wave, who had requested no future contact, or who were no longer on panel, were not recontacted. Any other eligibility criteria were applied during screening at the beginning of the survey.

6.7.1 Panel-owned Recontact sample: Interwave reminders, survey invitations, and survey reminders

- Panel firms' invitation and reminder email templates either followed their standard templates (different from ITC templates), or incorporated ITC's recommended phrasing (Appendix E).
- Panel firm details were provided in the email invitations and reminders, and in the survey, so that cohort panelists could contact the panel or to opt out.
- Contact information for the principal investigator and an ethics contact person was provided in the survey, and in some email invitations/reminders.
- ***Interwave reminders*** were sent to eligible sample before fieldwork started. The interwave reminders highlighted the importance of the survey and described the time commitment and remuneration value.
- ***Email survey invitations*** were sent to eligible panel-owned Recontact sample, in batches, beginning on Day 1 of fieldwork. Invitations described the time commitment and remuneration value for the 4CV5 Survey, and provided a personalized direct link to the survey.
- ***Email reminders*** were sent to all panelists who had been invited but not yet completed the survey per their respective panel firm's standard schedule, or at intervals determined in consultation with the ITC team, depending on survey activity.

6.7.2 Remuneration

- Remuneration was dependent on completing the 4CV5 Survey. See Table 8 for values.

6.7.3 Technical support

- Contact information for panelists' home panel firms were provided in recruitment emails and the 4CV5 Survey, for technical support and/or comments.
- If applicable, the home panel firm's toll-free number for was also provided in recruitment emails and the 4CV5 Survey, and a staff person was designated to check messages on the toll-free number on a daily basis during fieldwork.
- Survey firms were responsible to respond within 48 hours.

6.8 Fieldwork Procedures - CA, EN, and US Panel-owned Replenishment sample

- To account for attrition and maintain the target sample size in each country, new respondents were recruited at 4CV5.
- ITC DMC provided each survey firm with updated replenishment targets, at appropriate intervals during the fieldwork period as the actual retention rate (of Recontact sample) became known.
- **Preliminary eligibility criteria** for new recruits were determined from the survey firms' (Leger, Ipsos PA, Rakuten) existing knowledge of the panelists:
  - Any panelist who had reported being a smoker or former smoker at any panel update in the past 24 months, i.e., met preliminary eligibility criteria for a smoker/recent former smoker.
  - Any panelist who had reported being a NVP/HTP user and who screened into the survey as using NVP/HTP at least weekly.
  - Any other eligibility criteria were discussed and determined in consultation with the ITC Team.
- The panel firms emailed panelists who met preliminary eligibility criteria and provided them with a personalized link to the 4CV5 Survey.
- The final screening and allocation to Replenishment sample target quotas took place within the survey.

6.8.1 Replenishment sampling quotas

See Appendix B for the replenishment sample quotas by country.

6.8.2 Panel-owned Replenishment sample: Survey invitations and reminders

- Panel firms' invitation and reminder email templates either followed their standard templates (different from ITC templates), or incorporated ITC's recommended phrasing (Appendix E).
- Panel firm details were provided in the email invitations and reminders, and in the survey, so that panelists could contact the panel or to opt out.
- Contact information for the principal investigator and an ethics contact person was provided in the survey, and in some email invitations/reminders.
- ***Email survey invitations*** were sent to panel-owned Replenishment sample who met preliminary criteria in batches until the Replenishment targets with the corresponding sample characteristics were attained. Invitations described the time commitment and remuneration value for the survey, and provided a personalized direct link to the survey.
- ***Email reminders*** were sent to all panelists who had been invited but not yet completed the survey per their respective panel firm's standard schedule, or at intervals determined in consultation with the ITC team, depending on survey activity.

6.8.3 Remuneration

- Remuneration was dependent on completing the 4CV5 Survey. See Table 8 for values.

6.8.4 Technical support

- Contact information for panelists' home panel firms were provided in recruitment emails and the survey, for technical support and/or comments.
- The home panel firm provided technical support contact information in recruitment emails and the survey.

6.9 Study Remuneration

- Participants were required to submit their survey before being provided remuneration.
- Remuneration varied by sample source (Table 8).

Table 8. Summary of 4CV5 remuneration by country and sample source

Country/Panel	4CV5 Survey Remuneration	Processing responsibility
AU		
'ITC-owned' respondents **(all Recontact)**	AU $40 e-gift card	RM
RM/Dynata panelists **(Recontact, Replenishment)**	**RM (Recontact, Replenishment)**: AU $40 e-gift card **Dynata (Recontact, Replenishment)**: AU $10 worth of points	RM Dynata
CA		
'ITC-owned' respondents **(all Recontact)**	Cheque* for CA$28 ($25 + $3 bonus)	UW admin
Leger panelists **(Recontact, Replenishment)**	**Recontact:** CA $30 ($25 + $5 bonus) worth of points **Replenishment:** CA $25 worth of points	Leger
EN		
'ITC-owned' respondents **(all Recontact)**	£18 (£16 + £2 bonus) Amazon e-gift card	Logistic Solutions
Ipsos UK **(all Recontact)**	£18 worth of points	Ipsos UK
Rakuten **(Recontact, Replenishment)**	**Replenishment:** Standard credit for a survey of this length + additional bonus of £2.50 worth of points, with promise of an additional bonus of up to £15 worth of points after they complete the next wave's survey. **Recontact:** Standard credit for a survey of this length + £15 bonus.	Rakuten, Dynata
US		
'ITC-owned' respondents **(all Recontact)**	Check* for US $28 ($25 + $3 bonus)	UW admin
Ipsos PA panelists **(Recontact, Replenishment)**	**Recontact:** 13,000 (10,000 + 3,000 bonus) points **Replenishment:** 13,000 (10,000 + 3,000 bonus) points	Ipsos PA
Ipsos UK **(all Recontact)**	US $28 ($25 + $3 bonus) worth of points	Ipsos UK

** An Amazon e-gift card could be provided if requested by respondent and authorization given to input their email on a third party website.

7 Targets and Outcomes

This chapter presents summaries of recruitment numbers:

- Total completed surveys and cases removed (Table 9)
- Target vs. realized sample sizes (Table 10)
- Sample sizes and retention rates, by source (Table 11)
- Outcome rates by country, for new recruits (Table 12)
- Survey length by user type and country (Table 13)

Table 9. 4CV5 sample: All completed surveys vs. cases in the "Full Sample Dataset"* vs. cases in the "Main Sample Dataset"**. Note: Categories of cases removed

	AU	CA	EN	US	Total
Raw number of completed cases					
All completed surveys	1502	2456	2301	2201	8460
Cases removed to obtain the Full Sample Dataset* *Note: Categories are not mutually exclusive.*					
Speeders	7	27	12	25	71
Smoking or other status conflict between waves	0	1	0	5	6
Do not qualify for recruitment requirement	0	1	0	0	1
Did not hit submit button at end of survey	0	0	2	2	4
FULL SAMPLE DATASET					
Cases included in **Full Sample Dataset***	1495	2427	2275	2169	8378
Cases removed to obtain Main Sample Dataset* *Note: Categories are not mutually exclusive.*					
Quit over 5 years (and no NVP use)	14	58	12	75	159
MAIN SAMPLE DATASET					
Cases included in **Main Sample Dataset****	1481	2369	2275	2094	8219

***Full Sample Dataset =** completed cases excluding cases listed in the table.

****Main Sample Dataset =** full data excluding quitters quit >5 years.

Table 10. 4CV5 target sample sizes and (realized sample sizes) in the Full Sample Dataset*, by country

Subsample	CA	EN	US	AU
Main survey				
Smokers/former smokers aged 18+	1650 (1632)	1800 (1655)	1650(1628)	1500 (1495)
At-least-weekly NVP users aged 18+	802 (795)	500 (492)	500 (541)	
Total	2452 (2427)	2300 (2287)	2150 (2169)	1500 (1495)

* Full sample = Final realized 4CV5 sample. Excluded cases are listed in Table 9.

Table 11. Sources, Full Sample Dataset sample size [*,**], and retention rates for 4CV5 respondents

Country	Subsample/ user type	Sources	Eligible sample from 4CV43	Retention (%)	Recontact (n)	Replenishment (n)	Total (n)
AU	Smokers/Former Smokers 18+ ***	ITC-owned	82	54.9	45	0	45
		Roy Morgan	1429	46.7	667	783	1450
		AU Sub-Total	***1521***	***46.8***	***712***	***783***	***1495***
CA	Smokers/Former Smokers 25+	ITC-owned****	159	71.7	114	0	114
		Leger Marketing	1459	60.4	881	567	1448
	Smokers/Former Smokers 18-24	Leger Marketing	150	18.6	26	44	70
	NVP, HTP, snus, TF-ONP users 18+	Leger Marketing	416	53.4	222	573	795
		CA Sub-Total	***2184***	***56.9***	***1243***	***1184***	***2427***
EN	Smokers/Former Smokers 25+	ITC-owned	109	80.7	88	0	88
		Ipsos UK	386	47.2	182	0	182
		Rakuten	1178	9.8	115	1270	1385
	Smokers/Former Smokers 18-24	ITC-owned	0	0	0	0	0
		Ipsos UK	0	0	0	0	0
		Rakuten	92	4.3	4	136	140
	NVP, HTP, snus, TF-ONP users 18+	Ipsos UK	87	52.9	46	0	46
		Rakuten	423	1.7	7	439	446
		EN Sub-Total	***2312***	***19.1***	***442***	***1845***	***2287***
US	Smokers/Former Smokers 25+	ITC-owned****	66	75.8	50	0	50
		Ipsos PA	1519	69.0	1049	468	1517
		Ipsos UK	16	56.3	9	0	9
		ITC-owned (prev. opt-in)	51	82.4	42	0	42
	Smokers/Former Smokers 18-24	ITC-owned	0	0	0	0	0
		US opt-in	0	0	0	0	0
		Ipsos PA	12	25	3	6	9
		ITC-owned (prev. opt-in)	10	10	1		1
	NVP, HTP, snus, TF-ONP users 18+	Ipsos PA	492	71.1	350	180	530
		US opt-in	0	0	0	0	0
		ITC-owned (prev. opt-in)	2	50	1	0	1
		Ipsos UK	16	62.5	10	0	10
		US Sub-Total	***2184***	***69.4***	***1515***	***654***	***2169***
Full Sample Total			**8201**	**47.7%**	**3912**	**4466**	**8378**

*Full Sample Dataset = Realized 4CV5 sample (excludes low quality records, but includes >5year quitters). See Table 9 for the Main Sample Dataset (for which >5year quitters have been excluded).

*** In Australia, there was no split quota for smokers, recent former smokers and NVP users. Although NVP users were not specifically targeted, they were kept in the sample if recruited.

Table 12. Response rates and cooperation rates for new recruits at 4CV5, by country.

	Australia		England		Canada		US	
	Freq	%	Freq	%	Freq	%	Freq	%
A – Interviewed								
Total (interviewed)	788	39.9	1,856	3.8	1,203	5.9	670	10
B – Eligible, but not interviewed								
Refusal/breaks off	6	0.3	124	0.3	87	0.4	20	0
Other	0	0.0	0	0.0	0	0.0	0	0
Total (eligible but not interviewed)	6	0.3	124	0.3	87	0.4	20	0
C – Unknown if eligibility (not interviewed)								
Estimated number of eligible and quota not full*1	777	39.4	12,262	25.4	4,010	19.8	165	2
Estimated number of not eligible or quota full*2	204	10.3	30,305	62.8	11,476	56.7	1,766	25
Total (unknown if eligible)	981	49.7	42,567	88.3	15,486	76.5	1,931	28
D – Not eligible								
Out of sample	0	0.0	0	0.0	0	0.0	0	0
Respondent was found not to be eligible by SCSRU	56	2.8	2,812	5.8	3,381	16.7	3,840	55
Respondent was found not to be eligible by survey firm	0	0.0	0	0.0	0	0.0	0	0
Quota full	143	7.2	860	1.8	86	0.4	541	8
Other	0	0.0	0	0.0	0	0.0	0	0
Total (not eligible)	199	10.1	3,672	7.6	3,467	17.1	4,381	63
Total sample with final disposition	1,974	100	48,219	100	20,243	100	7,002	100
Estimated eligibility rate*3	93.4%		41.3%		27.6%		15.2%	
Estimated proportion for which quota was full*4	15.3%		30.3%		6.3%		43.9%	
Response rate*5	50.2%		13.0%		22.7%		78.4%	
Cooperation rate*6	99.2%		93.7%		93.3%		97.1%	

Notes:

*1 Estimated number of respondents that would have been eligible and for which the corresponding quota would not have been full

Formula: row 22 x row 34 x (1 - row 35), rounded to the nearest integer

*2 Formula: row 22 - row 20

*3 Estimated proportion of individuals that were found to be eligible

Formula: 1 - row 26 / (row 12 + row 17 + row 26)

*4 Estimated proportion of individuals that were terminated because the corresponding quota was full

Formula: row 28 / (row 12 + row 17 + row 28)

*5 The response rate is the proportion of eligible respondents who complete the survey; i.e., the number of eligible respondents who completed the survey divided by the estimated number of eligible respondents that were selected/contacted. In other words, the response rate accounts for the fact that numerous individuals could not be contacted or screened for eligibility, whereas the cooperation rate does not.

Formula: row 12 / (row 12 + row 17 + row 20)

Different survey firms used different screening process before sending respondents to the 4CV Survey. Consequently, the estimated eligibility rates (see row 34) differ greatly from one firm/country to another. Because of this (and a few other lesser reasons), the response rates are not comparable across countries/firms.

*6 The cooperation rate is the proportion of eligible respondents (i.e., those who have completed all eligibility questions and have been found to be eligible) who complete the survey.

Formula: row 12 / (row 12 + row 15)

The above formula for the cooperation rate is the same as the AAPOR COOP4 formula; see https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx

Contrary to the response rates, the cooperation rate are comparable across countries.

Table 13. 4CV5 survey length in minutes, by user type and by country.*

	AU		CA		EN		US		Overall	
	n	**median**	**n**	**median**	**n**	**median**	**n**	**median**	**n**	**median**
Smokers only	731	47.3	925	44.7	773	39.4	1013	45.5	**3442**	**44.4**
NVP users only	81	45.9	158	37.3	106	37.1	216	41.2	**561**	**39.8**
Concurrent users	378	59.6	848	45.9	1140	47.9	354	52.2	**2720**	**49.3**
Quitters	305	45.1	496	38.1	268	39.4	586	39.1	**1655**	**40.0**
Overall	**1495**	**49.0**	**2427**	**43.2**	**2287**	**42.8**	**2169**	**43.7**	--	--

* The questionnaire differed somewhat from country to country.

8 ANDS Image Upload Sub-study

There were two sub-studies at 4CV5

1) ANDS Image upload sup-study: Participants from all countries who used one or more of the following products at least weekly – NVPs, HTPs, snus, or TF-ONPs -- and who completed their main survey on a tablet or smart phone were invited the ANDs Image Upload Sub-study.
2) Experimental Tobacco Marketplace (ETM): In Canada, England, and the US, an Experimental Tobacco Marketplace (ETM) Sub-study recruited a limited number of respondents from the 4CV3 Survey.

8.1 ANDs Image Upload Sub-study

For the 4CV5 Survey, we recruited cigarette smokers/recent former smokers and users of 4 specific classes of ANDs. The 4CV5 Screening Section describes the physical characteristics of each of the five product types – cigarettes, NVPs, HTPs, snus (contrasted against other oral tobacco products), and TF-ONPs – provide country-specific examples and images of each product type. However, for some respondents, the distinction between vaping products and heated tobacco products may not be entirely clear.

Thus, the purpose of the ANDs Image Upload Sub-study is to validate if respondents accurately self-classify as NVP users compared to HTP users, or accurately as snus or tobacco-free oral tobacco users. Self-reported use in the survey is validated (i.e., classified as a 'match' vs. 'mismatch') against images of their products provided by survey respondents as well as against the brand information provided by respondents' 4CV5 survey responses.

8.1.1 ANDs Image upload Sub-study sample and eligibility criteria

At the end of the 4CV5 Survey, respondents who used NVPs, HTPs, snus, or TF-ONPs at least weekly and who had completed the survey on a tablet or smart phone were asked to upload a picture of their device and refill product (e.g., e-liquid, or IQOS heatsticks, etc.). If a respondent reported using more than one of NVPs, HTPs, snus, or tobacco-free ONPs at least weekly (and met the criteria for tablet/smart phone) they were asked to provide a picture of each product separately.

8.1.2 ANDs Image upload Sub-study sample procedures

1) Eligible sample were shown the survey variables below.

Varname	4CV5 Question
AI800	**Ask if EC309v=10-32 and device type=mobile or tablet.** Please upload a picture of [all of the vaping devices (with their e-liquids) that you currently use (EC309v=10-20)/ the vaping device and e-liquid that you used last (EC309v=31-32)]. In a single photo, please: * Include [all currently-used (EC309v=10-20)/ the last-used (EC309v=31-32)] vaping device(s) and e-liquid(s), and * Ensure brand name(s) are visible. Are you willing to upload a photo of your device(s) and e-liquid(s)? 1 Yes 2 No 8 Refused *Maximum upload size is 5 MB.*
AI801	**Ask if AI800=1.** Please tell us anything we should know about the picture you have uploaded. [open text field] *If you don't have any comments, please type in 'NONE'.*
AI805	**Ask if HN309v=10-32 and device type=mobile or tablet.**

	EC309v=10-32: Earlier, you told us that you use both vaping devices and heated tobacco products… **All**: Please upload a picture of [all of the heated tobacco devices that you currently use including the tobacco sticks/pods (HN309v=10-20)/ the heated tobacco device that you used last including its tobacco sticks/pods (HN309v=31-32)]. In a single photo, please: * Show [all of the heated tobacco devices that you currently use including the tobacco sticks/pods (HN309v=10-20)/ the heated tobacco device that you used last including its tobacco sticks/pods (HN309v=31-32)], and * Ensure brand name(s) are visible Are you willing to upload a photo of your device(s) and tobacco sticks/ pods/ capsules? 1 Yes 2 No 8 Refused *Maximum upload size is 5 MB.*
AI806	Ask if AI805=1. Please tell us anything we should know about the picture you have uploaded. [open text field] *If you don't have any comments, please type in 'NONE'.*
AI820	**Ask if NC006=1-3 and device type=mobile or tablet.** Please upload a picture of all of the snus product(s) that you currently use. In a single photo, please: * Show all of the snus product(s) and * Ensure brand name(s) are visible Are you willing to upload a photo of your snus product? 1 Yes 2 No 8 Refused *Maximum upload size is 5 MB.*
AI824	**Ask if AI820=1.** Please tell us anything we should know about the picture you have uploaded. [open text field] *If you don't have any comments, please type in 'NONE'.*
AI827	**Ask if BI038=1-3 and device type=mobile or tablet.** Please upload a picture of all of the tobacco-free oral nicotine product(s) that you currently use. In a single photo, please: * Show all of the tobacco-free oral nicotine product(s) and * Ensure brand name(s) are visible Are you willing to upload a photo of your tobacco-free oral nicotine product(s)? 1 Yes 2 No 8 Refused *Maximum upload size is 5 MB.*
AI831	**Ask if AI827=1.** Please tell us anything we should know about the picture you have uploaded. [open text field] *If you don't have any comments, please type in 'NONE'.*

2) Images were encrypted and uploaded to a secure server hosted at the University of Waterloo.
3) After fieldwork was closed, the research team downloaded the images and coded each image using a unique identifier that was linked via a master list to the 4CV5 record number.
4) Images were independently classified as NVP, HTP, snus, TF-ONP, or other by independent coders. The classifications were subsequently compared with 4CV5 user status, based on 4CV5 responses.

8.2 Experimental Tobacco Marketplace (ETM) Sub-study

At 4CV5, upon completion of the main survey, a subsample of respondents was invited to participate in an additional sub-study, the Experimental Tobacco Marketplace (ETM). This sub-study was led by researchers from Virginia Tech (VT) and Medical University of South Carolina (MUSC).

The ETM is an online store that displays photos, prices, and information for each of several available products. The ETM thus provides a method to estimate, prior to implementation, the effects of new products or policies on purchasing behaviors across various products in a complex tobacco marketplace. The ETM allows participants to make hypothetical purchases. A further methodological development of this approach, the Illegal Tobacco Marketplace (iETM) include marketplace choice or preference as a way to assess the potential of illicit trades under different policy restrictions.

The purpose of this sub-study was to examine the effects of implementing a vaping ban and a flavored vaping ban on the probability of purchasing vaping products in an illegal experimental tobacco marketplace as the price of conventional cigarettes increased in a legal experimental tobacco marketplace. The ETM sub-study was conducted using samples from CA, EN, and the US (countries with different policy environments) to ascertain whether the local regulatory conditions would modulate the observations. The target sample in each country was 150 overall, including 75 smokers only and 75 daily smoker/weekly vaper concurrent users.

8.2.1 ETM sample, eligibility criteria, and procedures

- Eligible panel-owned respondents for the ETM sub-study were recruited in CA, EN, and US from the following firms only: Leger (CA), Rakuten (EN), Ipsos PA (US). Due to resources, the ETM sub-study was offered in the English language only. For Canada, due to language rights this meant only linking English-speaking respondents to the ETM.
- Eligible panelists who agreed to participate in the ETM sub-study were redirected to the ETM Survey, which was programmed in Qualtrics and hosted on a VT server. The recruitment script included "*Please note: Although we will try to only invite 150 people, if the survey is very popular, it may fill up quickly and there is a small chance your user group may be filled.*"
- Upon completion of the ETM survey, participants were then redirected back to their respective home panel firm's website via a URL so a records of their ETM survey completion could be enumerated, and the home panel firm would then process each participant's remuneration.
- Once the sample target for the respective country and user type was achieved, no further sample was recruited to the ETM sub-study for that target. The VT team monitored targets and alerted the ITC/panel firm teams so that sample invitations would be slowed and/or closely monitored.

Respondents were identified from the main survey using the following ETM eligibility criteria:

a) 4CV5 panel-owned respondents aged 18y or older who were **smokers only**, defined as:
   - daily cigarette smoker (FR309v=10) for at least 2 months
   - not NVP user (EC309v=40-70)
   - smoked factory-made cigarettes only, more than, or about the same amount as roll-your-own (FR330=1-3)

b) 4CV5 panel-owned respondents aged 21y or older who were **dual daily users**, defined as:
   - daily cigarette smoker (FR309v=10) for at least two months
   - weekly e-cigarette user (EC309v=10 or 20) for at least two months
   - smoked factory-made cigarettes only, more than, or about the same amount as roll-your-own (FR330=1-3)

8.2.2 Completed Respondents by User Group and Country

The number of 4CV5 Survey participants who were invited to the ETM, the number who agreed to participate in the ETM, and the number that completed the ETM are shown below in Table 14. The final analyzed sample included in the ETM data set is shown in Table 15; cases were excluded if their self-reported data were inconsistent with 4CV5 Survey data.

Table 14. Obtained sample sizes of participants completing the survey for each of three user groups in three countries.

Country	Shown ETM invitation	Agreed to participate	Completes	
			Smokers Only	Daily smoker/ weekly vapers
CA	316	202	95	78
EN	287	229	71	78
US	271	223	101	28

9 Quality Control and Uniformity

9.1 Incongruent cases and requests for withdrawal

- Any cases in which a respondent's identity was determined to be incorrect, or in which data were incongruent across waves (i.e., mismatched user status and/or duplicate cases), or in which a respondent requested their data be withdrawn, were subsequently removed from the full data set. Any ineligible cases (e.g., someone filled out survey in place of someone else) were also removed from the full data set.
- The number of incongruent cases removed from the dataset is provided in Table 9.

9.2 Identification and removal of 'satisficers/speeders' from the dataset

- The number of 'speeders/satisficers' removed from the dataset is provided in Table 9.
- Since Wave 1, two criteria were used to identify for poor data quality:
  - Seconds per question (secperQ), and
  - % of responses that were either Refused or Don't Know (%RDK).
    - Very extreme values occurred for both of these variables: times of less than 1.1 seconds per question, which by published estimates is not sufficient for even reading the question, and RDK responses for more than 93% of the questions completed.
- The respondents were split by country and user groups. Frequency distributions by country and user groups were used to create precise cut-off values.
- 'Suspiciousness points' were assigned, based on the cut-offs for the respondent's country and user group, and all respondents had points assigned. Respondents were removed from the dataset if they scored too many points.
- Points were assigned on these bases:
  1) SecperQ -- The calculated value was time taken to complete the survey divided by the number of questions answered by the respondent. Very short times suggest poor data quality.
     - 5 points were assigned if the respondent's value was lower than $0.5*10^{th}$ percentile.
     - 3 points were assigned if the respondent's value fell between $0.5*10^{th}$ percentile and 0.75*10th percentile.
     - 2 points were assigned if the respondent's value fell between $0.75*10^{th}$ percentile and the 10^{th} percentile.
  2) %RDK -- The calculated value was the number of questions answered with either Refused or Don't know, divided by the number of questions answered by the respondent. Large numbers represent poor data quality.
     - 3 points were assigned if the respondent's value was higher or equal to $2*90^{th}$ percentile.
     - 2 points were assigned if the respondent's value fell between $2*90^{th}$ percentile and $1.33*90^{th}$ percentile.
     - 1 point was assigned if the respondent's value fell between $1.33*90^{th}$ percentile and the 90^{th} percentile.
- Points results: The range of possible points was 0-8. 81.55% of the respondents had 0 points.
- Cut-off for dropping respondents: Respondents who had more than 4 points were removed from the final data. In this setting, those who had extreme values in SecperQ (lower than $0.5*10^{th}$ percentile) were out irrespective of their %RDK value. This resulted in dropping a total of 71 respondents from the final data: Canada had 27 dropped respondents, followed by 25 in US, 12 in England, and 7 in AU.
- Comparison between respondents who were dropped and those who were kept in the dataset indicated that dropped respondents were:
  - more likely to be in the 18-24 and 25-39 age groups, while 55+ group had the least dropped respondents;
  - more likely to be male than female.

Appendix A: Sample sizes, retention rates of 4CV1-5

Table A.1. Sample sizes and retention rates at 4CV1 to 4CV5*

4CV1				
	ITC-owned	Panel-owned		
	Recontact n *(% retention)*	Recontact n *(% retention)*	Replenishment n	Total n *(% retention)*
Australia	276 *(48.9)*	239 *(38.2)*	989	1504 *(43.1)*
Canada	524 *(39.1)*	137 *(54.2)*	3167	3828 *(41.6)*
England	304 *(28.4)*	0	4070	4374 *(35.7)*
United States	296 *(27.8)*	1161 *(54.2)*	1355	2812 *(44.2)*
Total	1400 *(34.7)*	1537 *(50.9)*	9581	12518 *(41.6)*
4CV2				
	ITC-owned	Panel-owned		
	Recontact n *(% retention)*	Recontact n *(% retention)*	Replenishment n	Total n *(% retention)*
Australia	199 *(72.1)*	654 *(53.3)*	662	1515 *(56.7)*
Canada	331 *(63.2)*	1537 *(46.5)*	1915	3783 *(48.8)*
England	206 *(67.8)*	1525 *(37.5)*	3117	4848 *(39.6)*
United States	236** *(54.1***)*	969 *(38.5)*	1643	2848 *(42.9)*
Total	972 *(69.4)*	4685 *(42.1)*	7337	12994 *(45.2)*
4CV3				
	ITC-owned	Panel-owned		
	Recontact n *(% retention)*	Recontact n *(% retention)*	Replenishment n	Total n *(% retention)*
Australia	137 *(68.8)*	738 *(56.1)*	616	1491 *(57.8)*
Canada	224 *(67.7)*	1437 *(41.6)*	2027	3688 *(43.9)*
England	135 *(65.5)*	1451 *(31.3)*	2314	3900 *(32.7)*
United States	215**(91.5***)*	1149 *(44.0)*	1164	2528 *(47.9)*
Total	711 *(73.1)*	4775 *(39.7)*	6121	11607 *(42.2)*
4CV4				
	ITC-owned	Panel-owned		
	Recontact n *(% retention)*	Recontact n *(% retention)*	Replenishment n	Total n *(% retention)*
Australia	*82 (59.9)*	*686 (50.7)*	*753*	*1521 (51.5)*
Canada	*159 (71.0)*	*1298 (37.5)*	*727*	*2184 (39.5)*
England	*109 (80.7)*	*946 (25.1)*	*1257*	*2312 (27.1)*
United States	*66 (30.7)*	*1204 (52.1)*	*914*	2184 (50.2)
Total	*416 (58.5)*	*4134 (37.9)*	*3651*	*8201 (39.2)*
4CV5				
	ITC-owned	Panel-owned		
	Recontact n *(% retention)*	Recontact n *(% retention)*	Replenishment n	Total n *(% retention)*
Australia	45 *(54.9%)*	667 *(46.7%)*	783	1495 *(46.8%)*

Canada	114 *(71.7%)*	1129 *(55.8%)*	1184	2427 *(56. 9%)*
England	88 *(80.7%)*	354 *(16.1%)*	1845	2287 *(19.1%)*
United States	94 *(72.9%)*	1463 *(71.2%)*	654	2169 *(69.4%)*
Total	341 *(71.2%)*	3571 *(46.1%)*	4466	8378 *(47.7%)*

C= Recontact; P= Replenishment

* Using 4CV1 main sample; and full sample for 4CV2-5.

**Includes 121 new ITC-owned originally recruited as US opt-in at 4CV1 and 4CV2.

*** With new ITC-owned, retention rate = (94+121)/(236) = 91.5%. For new ITC-owned only, retention rate = (94/160) = 58.8%.

APPENDIX B: Replenishment sample targets, by country

In order to ensure the Replenishment Sample was representative, the research team determined approximate quota targets for Replenishment subsample groups as necessary, depending on the quality of the sample provided by panel firms. Quota targets were based on one or more of: user type, age, sex, and region. The quotas used in each country are described below.

Appendix B.1: Australia replenishment sample targets (Roy Morgan, RM)

- Replenishment sample were recruited by RM. RM subcontracted Dynata (previously named Survey Sampling International, SSI).
- The DMC established AU quota targets using **age** and **sex** criteria. Targets were proportional to stratum sizes based on AU census data (Table B.1).
- Within each quota target, the AU replenishment sample was recruited from a **probability-based panel** (RM Single Source) and a **non-probability-based panel** (via RM's panel partner, Dynata).
- UW monitored survey response in the cohort (recontact) sample (from ITC-owned, RM, and RM's panel partner) and adjusted the replenishment sample quotas (for RM and RM's panel partner) during fieldwork to achieve the final representative sample.

Table B.1. Summary of 4CV5 Australia replenishment sample user groups and quota criteria

4CV5 AU - Replenishment sample user groups	**Quota target criteria**			
	Age	**Sex**	**Region**	**Language**
Smokers, Recent Former Smokers, NVP Users	Yes	Yes	No	English only

** In Australia only the target sample was for smokers/recent former smokers/NVP users as a group (i.e., there were no subtargets based on user type).*

Figure B.1. Initial Replenishment Sample quota targets for Australia.

Note: Targets were updated periodically during fieldwork as appropriate to reflect the actual Recontact/ Replenishment completes.

Quotas for Australia
ITC 4CV Survey Wave 5
Created by C. Boudreau on Mar 4, 2024
Last updated by C. Boudreau on Nov 27, 2024

Notes:
1) All quotas in this spreadsheet are for individuals that are to be recruited at wave 5; individuals recruited at prior waves that will be recontacted at wave 5 do not count towards the quotas in this spreadsheet
2) All quotas must be meet or slightly exceeded by the end of fieldwork
3) The quotas to be programmed are in orange; the other numbers in the table are totals that do not need to be programmed
4) Cells in gray are meant to be easily modified (just type in a new number); all other cells/calculations are automated and thus those cells shouldn't be modified

Quotas for smokers/recent quitters ages 18 & older (regardless of e-cig/HTP/snus usage)

		Estimated # indvs *1		Quotas	
		N	%	Original	Revised
Male	18-24	278,295	8.0%	53	64
	25-39	683,180	19.6%	131	158
	40-54	577,857	16.6%	111	133
	55+	414,095	11.9%	79	96
	Total	1,953,427	56.1%	374	345
Female	18-24	167,313	4.8%	32	39
	25-39	538,932	15.5%	103	124
	40-54	457,943	13.2%	88	106
	55+	362,290	10.4%	69	84
	Total	1,526,478	43.9%	292	353
Total		3,479,905	100.0%	666	803
				*2	*3

Notes:
*1 Estimated numbers of smokers, recent quitters, e-cigarettes, HTP and/or snus users age 18 & older
These estimates are from the 2019 National Drug Strategy Household Survey (NDHSH)
*2 This (initial) estimate of 666 respondents to be recruited at wave 5 was computed by M. Thompson as part of the P01 grant application
*3 This revised estimate of 803 respondents to be recruited at wave 5 is based on a 46.5% retention rate (as of Nov 27, 2024), and a desired overall sample of 1500 respondents to be interviewed at wave 5; 1500 - 1497 x 0.465 = 803

Appendix B.2: Canada replenishment sample targets (Leger)

- The DMC established CA smoker/recent former smoker and NVP/HTP/snus/TF-ONP user quota targets using **sex, age, region,** and **language** criteria. Targets were proportional to stratum sizes based on CA census data.
- *Important: Users who met criteria for both the smoker/recent former smoker and NVP/HTP/snus/TF-ONP user subsamples were first counted towards the appropriate NVP/HTP/snus/TF-ONP sample target. Once the NVP/HTP/snus/TF-ONP sample targets were met, subsequent concurrent users were counted towards the appropriate smoker targets.*
- The sampling **regions** were**:**
    1) Maritimes (incl. Newfoundland and Labrador, Prince Edward Island, Nova Scotia, and New Brunswick)
    2) Quebec
    3) Ontario
    4) Prairies (incl. Manitoba, Saskatchewan, Alberta)
    5) British Columbia
- If the quota value was very small (e.g., for NVP/HTP/snus/TF-ONP users in Maritimes provinces), they were presented as one value for the aggregated regions.
- Within each quota target, the CA replenishment sample was recruited from a **probability-based panel** (via Leger).
- The survey firm and UW monitored survey response in the cohort (recontact) sample (from ITC-owned and Leger) and adjusted the replenishment sample quota targets during fieldwork to achieve the final representative sample.

Table B.2. Summary of 4CV5 Canada replenishment sample user groups and quota criteria

4CV5 CA - Replenishment sample user groups	**Quota criteria**			
	Age	**Sex**	**Region**	**Language**
Smokers, Recent Former Smokers	Yes	Yes	Yes	No specific targets.
NVP, HTP, snus, TF-ONP users	Yes	Yes	Yes	No specific targets.

Figure B.2. Initial 4CV5 Replenishment sample targets for Canada.

Note: Targets were updated periodically during fieldwork as appropriate to reflect the actual Recontact/ Replenishment completes.

Quotas for Canada

ITC 4CV Survey Wave 5

Created by C. Boudreau on Mar 4, 2024

Last updated by C. Boudreau on Mar 8, 2024

Notes:

1) All quotas in this spreadsheet are for individuals that are to be recruited at wave 5; individuals recruited at prior waves that will be recontacted at wave 5 do not count towards the quotas in this spreadsheet
2) All quotas must be meet or slightly exceeded by the end of fieldwork
3) The quotas to be programmed are in orange; the other numbers in the table are totals that do not need to be programmed
4) Cells in gray are meant to be easily modified (just type in a new number); all other cells/calculations are automated and thus those cells shouldn't be modified

Quotas for smokers/recent quitters ages 18 & older (regardless of e-cig/HTP/snus usage)

		Population Estimates*1					Prevalence*2					Estimated # of users*3						Quotas					
		Maritimes	Québec	Ontario	Prairies	BC	Maritimes	Québec	Ontario	Prairies	BC	Maritimes	Québec	Ontario	Prairies	BC	Total	Total	Maritimes	Québec	Ontario	Prairies	BC
Male	18-24	96,175	321,640	631,520	291,670	204,280	34.12	23.73	26.78	25.94	23.10	32,818	76,321	169,123	75,666	47,179	401,107	79	6	15	33	15	10
	25-44	276,290	1,094,980	1,875,690	949,605	681,080	27.48	17.36	19.31	16.50	14.42	75,923	190,075	362,198	156,729	98,214	883,139	175	15	38	72	31	19
	45 & older	594,410	1,939,035	3,060,540	1,330,205	1,124,435	19.05	13.99	14.93	13.45	13.72	113,227	271,298	456,926	178,846	154,256	1,174,553	231	22	53	90	35	31
	Total (Mal	966,875	3,355,655	5,567,750	2,571,480	2,009,795						221,968	537,694	988,247	411,241	299,649	2,458,799						
Female	18-24	89,385	308,605	586,290	270,745	192,400	15.79	14.71	15.67	14.80	22.80	14,115	45,393	91,875	40,064	43,875	235,322	47	13	9	18	8	9
	25-44	286,220	1,091,200	1,919,110	953,535	687,685	18.31	17.86	15.08	16.68	8.85	52,398	194,928	289,308	159,093	60,877	756,604	150		39	57	32	12
	45 & older	654,950	2,095,200	3,419,765	1,412,010	1,242,440	14.06	14.70	10.70	12.80	10.73	92,056	308,087	365,898	180,775	133,316	1,080,132	214	18	61	72	36	27
	Total (Fem	1,030,555	3,495,005	5,925,165	2,636,290	2,122,525						158,569	548,408	747,081	379,932	238,068	2,072,058						
Total		1,997,430	6,850,660	11,492,915	5,207,770	4,132,320						380,537	1,086,102	1,735,328	791,173	537,717	4,530,857	896	74	215	342	157	108
																		*4					

Quotas for e-cig/HTP/snus ages 18 & older

		Pop Estimates Canada	Prevalence*4	Estimated # of users*3 #	%	Quotas
Male	18-24	1,545,285	13.06	201,740	22.5%	152
	25-44	4,877,645	4.88	237,849	26.6%	180
	45 & older	8,048,625	1.27	102,057	11.4%	77
	Total (Mal	14,471,555		541,646	60.5%	409
Female	18-24	1,447,425	10.82	156,598	17.5%	118
	25-44	4,937,750	2.13	105,258	11.8%	80
	45 & older	8,824,365	1.03	91,138	10.2%	69
	Total (Fem	15,209,540		352,994	39.5%	267
Total		29,681,095		894,640	100.0%	676 *7

Notes:

*1 Population estimates from the 2021 Census; see spreadsheet "9810002201-eng-2" in this Excel file
*2 Esimated prevalence of smoking and/or using e-cig and/or quit within last 2 years
These estimates are from the 2019 Canadian Tobacco and Nicotine Survey (CTNS)
*3 Population estimates x esimated prevalences
*4 This (initial) estimate of 896 respondents to be recruited at wave 5 is based on the table from Anne (see below)
*5 Esimated prevalence of using e-cig
These estimates are from the 2019 Canadian Tobacco and Nicotine Survey (CTNS)
*6 This (initial) estimate of 676 respondents to be recruited at wave 5 is based on the table from Anne (see below)

Appendix B.3: 4CV5 Replenishment sample targets for England (Rakuten Insight and panel partners)

- The DMC established EN smokers/recent former smoker and NVP/HTP/snus/TF-ONP user quota targets using **sex, age,** and **region** criteria. Targets were proportional to stratum sizes based on EN census data.
- *Important: Users who met criteria for both the smoker/recent former smoker and NVP/HTP/snus/TF-ONP user subsamples were first counted towards the appropriate NVP/HTP/snus/TF-ONP sample target. Once the NVP/HTP/snus/TF-ONP sample targets were met, subsequent concurrent users were counted towards the appropriate smoker targets.*
- The sampling **regions** were**:**
    1) North and Yorkshire & the Humber
    2) Midlands
    3) London and East of England
    4) South
- Within each quota target, the EN replenishment sample was recruited from a **non-probability-based panel** (via Rakuten's panel partners).
- UW monitored survey response in the cohort (recontact) sample (from ITC-owned and Ipsos UK) and adjusted the replenishment sample quotas (for Rakuten) during fieldwork to achieve the final representative sample.

Table B.3. Summary of 4CV5 England replenishment sample user groups and quota criteria

4CV5 EN - Replenishment sample user groups	**Quota criteria**			
	Age	**Sex**	**Region**	**Language**
Smokers, Recent Former Smokers	Yes	Yes	Yes	English only
NVP,HTP, snus, TF-ONP users	Yes	Yes	No	English only

Figure B.3. Initial 4CV5 Replenishment sample targets for England

Note: Targets were updated periodically during fieldwork as appropriate to reflect the actual Recontact/ Replenishment completes).

Quotas for England (EN)
ITC 4CV Survey Wave 5
Created by C. Boudreau on Mar 4, 2024
Last updated by C. Boudreau on Sep 5, 2024

Notes:
1) All quotas in this spreadsheet are for individuals that are to be recruited at wave 5; individuals recruited at prior waves that will be recontacted at wave 5 do not count towards the quotas in this spreadsheet
2) All quotas must be meet or slightly exceeded by the end of fieldwork
3) The quotas to be programmed are in orange; the other numbers in the table are totals that do not need to be programmed
4) Cells in gray are meant to be easily modified (just type in a new number); all other cells/calculations are automated and thus those cells shouldn't be modified

Quotas for smokers/recent quitters ages 18 & older (regardless of e-cig/HTP/snus usage)

		Estimated # smokers *1	% Pop *2,*3				Quotas (initial)					Quotas (updated)				
			North, Yorkshire &	Midlands	London and East of England	South	Total	North, Yorkshire &	Midlands	London and East of England	South	Total	North, Yorkshire &	Midlands	London and East of England	South
Male	18-24	326,595	28.1%	19.6%	26.8%	25.6%	60	17	12	16	15	81	23	16	22	21
	25-34	699,424	26.2%	18.2%	32.3%	23.4%	128	33	23	41	31	174	46	32	56	41
	35-49	875,196	25.3%	17.7%	32.1%	25.0%	161	41	28	52	40	217	55	38	70	54
	50 & older	1,171,838	27.8%	19.5%	24.2%	28.5%	215	60	42	52	61	291	81	57	70	83
	Total (Males)	3,073,053					564					763				
Female	18-24	286,185	28.4%	19.5%	27.2%	24.9%	53	15	10	14	14	71	20	14	19	18
	25-34	596,076	26.1%	18.1%	32.6%	23.2%	109	28	20	36	25	148	39	27	48	34
	35-49	720,854	25.6%	17.9%	30.9%	25.7%	132	34	24	41	33	179	46	32	55	46
	50 & older	1,050,537	27.8%	19.3%	24.2%	28.7%	193	54	37	47	55	261	73	50	63	75
	Total (Females)	2,653,652					487					659				
Total		5,726,705					1052 *4					1422 *10				

Quotas for e-cig/HTP/snus ages 18 & older

		N *5	E-cigarettes			Quotas	
			Prevalence *6	Estimated # *7	%	initial	updated
Male	18-24	2,383,900	12.3	293,220	7.2%	26	32
	25-34	3,974,000	11.1	441,114	10.8%	39	49
	35-49	5,436,000	11.4	619,704	15.2%	55	68
	50-59	3,686,100	9.3	342,807	8.4%	30	38
	60 & older	7,057,200	9.2	649,262	16.0%	57	71
	Total (Males)	22,537,200		2,346,107		207	258
Female	18-24	2,253,420	18.6	419,136	10.3%	37	46
	25-34	3,748,900	9.1	341,150	8.4%	30	38
	35-49	5,502,700	7.5	412,703	10.1%	37	45
	50-59	3,811,400	7.3	278,232	6.8%	25	31
	60 & older	7,994,400	3.4	271,810	6.7%	24	30
	Total (Females)	23,310,820		1,723,031		153	190
Total		45,848,020		4,069,138	100.0%	360	448
						*8	*11

Notes:
*1 Estimated number of smokers as of 2024; see spreadsheet "Nb smk" in this Excel file
*2 Estimated percentage of population in each 4 geographic regions per gender/age group; see spreadsheet "Nb smk" in this Excel file
See other spreadsheets in this Excel file
*3 North, Yorkshire & Humber = North East, North West and Yorkshire & the Humber
Midlands = East Midlands and West Midlands
London and East of England = East of England and London
South = South East and South West
*4 This (initial) estimate of 1052 respondents to be recruited at wave 5 is based on the table from Anne (see below)
*5 Population estimates/projections for 2024
*6 Estimated prevalence of e-cigarettes use as of 2022; see spreadsheet "Table 1 - ecig" in this Excel file
*7 Estimated number of e-cigarettes users as of 2024; column D x column E
*8 This (initial) estimate of 360 respondents to be recruited at wave 5 is based on the table from Anne (see below)
*9 As of Sep 5, 86 of the expected 88 ITC-owned cohort respondents had been sucesfully recontacted at wave 5, but only 292 of the expected 660 18+ smokers/recent quitters and only 52 of the expected 140 18+ EC/HTP/snus/ONP users had been sucesfully recontacted at wave 5
*10 This (updated) estimate of 1422 respondents to be recruited at wave 5 is based on the achived recontact figures above (see *9), and a desired overall sample of 1800 (88+660+1052=1800) 18+ smokers/recent quitters to be interviewed at wave 5; 1800
*11 This (updated) estimate of 448 respondents to be recruited at wave 5 is based on the achived recontact figures above (see *9), and a desired overall sample of 500 (140+360=500) 18+ EC/HTP/snus/ONP users to be interviewed at wave 5; 500 - 52 =

Appendix B.4: 4CV5 Replenishment sample targets for the United States (Ipsos Public Affairs)

- Ipsos PA determined representative sampling procedures for sampling drawn from their high-quality **probability-based KnowledgePanel sample**.
- UW monitored survey response in the cohort sample (from ITC-owned, Ipsos PA, and Ipsos UK) and adjusted the replenishment sample targets for Ipsos' KnowledgePanel during fieldwork.
- *Important: Users who met criteria for both the smoker and NVP/HTP/snus/TF-ONP user subsamples were first counted towards the appropriate NVP/HTP/snus/TF-ONP quota. Once the NVP/HTP/snus/TF-ONP sample targets were met, subsequent users who met both criteria were counted towards the appropriate smoker targets.*

Table B.4. Summary of 4CV5 United States replenishment sample user groups and quota criteria

4CV5 US - Replenishment sample user groups	**Quota criteria**			
	Age	**Sex**	**Region**	**Language**
Smokers, Recent Former Smokers	Yes	Yes	No	English only
NVP/HTP/snus/TF-ONP Users	Yes	Yes	No	English only

Figure B.4. Initial 4CV5 Replenishment sample targets for the United States

Note: Targets were updated periodically during fieldwork as appropriate to reflect the actual Recontact/ Replenishment completes.

Quotas for US
ITC 4CV Survey Wave 5
Created by C. Boudreau on Mar 4, 2024
Last updated by C. Boudreau on Mar 8, 2024

Notes:
1) All quotas in this spreadsheet are for individuals that are to be recruited at wave 5; individuals recruited at prior waves that will be recontacted at wave 5 do not count towards the quotas in this spreadsheet
2) All quotas must be meet or slightly exceeded by the end of fieldwork
3) The quotas to be programmed are in orange; the other numbers in the table are totals that do not need to be programmed
4) Cells in gray are meant to be easily modified (just type in a new number); all other cells/calculations are automated and thus those cells shouldn't be modified

Quotas for smokers/recent quitters ages 18 & older (regardless of e-cig/HTP/snus usage)

		Estimated # indvs *1		Quotas				
		N	%	Initial	Revised (???)	Revised (???)	Revised (???)	Revised (???)
Male	18-24	3,068,234	6.8%	42				
	25-39	8,452,313	18.8%	115				
	40-54	7,098,710	15.8%	97				
	55+	7,222,989	16.1%	98				
	Total	25,842,246	57.6%	352				
Female	18-24	2,324,984	5.2%	32				
	25-39	5,152,982	11.5%	70				
	40-54	5,242,152	11.7%	71				
	55+	6,306,043	14.1%	86				
	Total	19,026,161	42.4%	259				
Total		44,868,407	100.0%	611				
				*2				

	Estimated # indvs *1		Quotas				
	N	%	Initial	Revised (???)	Revised (???)	Revised (???)	Revised (???)
Northeast	6,449,787	14.4%	88				
Midwest	10,492,213	23.4%	143				
South	19,436,469	43.3%	264				
West	8,489,939	18.9%	116				
Total	44,868,407	100.0%	611				

Quotas for e-cig/HTP/snus ages 18 & older

		Estimated # indvs *3		Quotas				
		N	%	Initial	Revised (???)	Revised (???)	Revised (???)	Revised (???)
Male	18-24	2,321,340	14.3%	28				
	25-39	4,290,859	26.4%	53				
	40-54	2,294,575	14.1%	28				
	55+	1,390,951	8.6%	17				
	Total	10,297,726	63.5%	126				
Female	18-24	1,760,524	10.8%	21				
	25-39	2,144,478	13.2%	26				
	40-54	1,237,066	7.6%	15				
	55+	788,054	4.9%	10				
	Total	5,930,121	36.5%	72				
Total		16,227,847	100.0%	198				
				*4				

Notes:
*1 Estimated numbers of smokers, recent quitters, e-cigarettes, HTP and/or snus users age 18 & older
These estimates are from the 2018 National Health Interview Survey (NHIS)
*2 This (initial) estimate of 611 respondents to be recruited at wave 5 is based on the table from Anne (see below)
*3 Estimated numbers of e-cigarettes, HTP and/or snus users age 18 & older
These estimates are from the 2018 National Health Interview Survey (NHIS)
*4 This (initial) estimate of 198 respondents to be recruited at wave 5 is based on the table from Anne (see below)

APPENDIX C: 4CV5 Fieldwork challenges

4CV5 Fieldwork challenges

1) Retention in England
   - We experience a much lower retention rate in England than in other countries, meaning a much higher reliance on fresh sample.
   - Rakuten has changed partner firms going forward and hope for much higher recontact rates in future waves.
2) ETM US Dual User targets
   - Ipsos PA was unable to provide the targeted 75 dual user respondents for the US version of the ETM.
   - Virginia Tech altered the targets to allow for a greater number of cigarette only respondents.
   - Virginia Tech reduced the screening of dual user targets from daily users of both cigarettes and vapes to weekly user of vapes who smokes at least daily.
   - Canadian targets for dual users were increased in the hope that it could make up for some of the US shortfall.
3) Age 18-24 Recruitment in Australia
   - RM was unable to provide the targeted number of 18-24 year-old respondents.
   - Snowball methodology was utilized to allow for an increased number of 18-24-year-olds.

Issues resolved from the previous wave(s)

1) Harder-to-recruit younger adults aged 18-24 were no longer oversampled as of Wave 4 (this was done in Waves 1-3).

2) The project management team successfully coordinated Replenishment sample target updates (across panel firm, and accounting for dynamic Recontact completes) at this wave. There was a slight overshoot, but budget savings offset this, and the benefit of a slightly more robust sample was deemed appropriate by the leadership team.

APPENDIX D: Terminology guidelines for publications

The standard terminology presented below was determined in consensus by the 4CV investigator team and finalized on October 12, 2018. It provides standard terms for use in the P01 Supplement and other manuscripts using 4CV data. Recommended main terms are in **bold;** optional qualifiers are in **square brackets**. Qualifiers may be omitted after first use if this intention is stated. For complete clarity, all terms for products, product classes, and user groups should be defined in a paper when first used. Where an intended classification does not match the definition for one of the recommended main terms, another explicit term should be used and carefully defined.

*Suggested people-first language is provided where relevant in the table below. Other suggested terms can be found at https://tobaccocontrol.bmj.com/pages/people-first-language-policy.

Table D.1: Terminology guidelines for publications

Concept (products)	Main term	Definitions	Comments
Cigarette	**Cigarette [roll-your-own or RYO, factory made or FM]**	Traditional cigarette made with tobacco, roll-your-own or factory made.	The term is unlikely to require the use of the "conventional" or "traditional" qualifier as the term "e-cigarette" will no longer be used routinely in papers (except where explicitly describing questions which include it, or in defining a NVP).
Device to deliver substance in vapour or aerosol form for inhaling.	**Vaping device**	A device, sometimes called a vaporizer or personal vaporizer, designed to produce from a liquid (called a vaping liquid) a vapour or aerosol for inhaling. This definition includes disposable e-cigarettes, rechargeable cartridges or pods which liquids, tank devices; they can be open or closed systems; the liquid may or may not contain nicotine.	Preferred term for the devices. Please do not use if you intend to include heated tobacco products. Alternative not preferred is "vaporizer". Do not use the term "e-cigarette" or the like except in the context of defining vaping device or NVPs.
	Nicotine vaping product (NVP)	A vaping device or vaporizer delivering nicotine in the vapour or aerosol.	The word "product" here and later is being used to refer to the combination of the device and the substance(s) it contains; i.e. where the focus is on the functional use, rather than on the physical device. Only use where the products are known to contain nicotine.
	Cigalike	A vaping device designed to resemble a cigarette.	Use only if necessary. Use another term such as "tobacco stick" or "heeted cigarette" for something like a Heets stick.

Concept (products)	Main term	Definitions	Comments
	Electronic nicotine delivery system (ENDS)	An electronic or battery operated device designed to deliver an inhaled dose of nicotine or other substances; not for Nicotine Replacement Therapy.	ENDS could refer to NVPs or HTPs – but FCA says not. (See second Appendix.) Avoid using this term if possible.
	Heated tobacco product (HTP)	Products that heats tobacco but to a lower temperature than needed for combustion, to deliver an aerosol for inhaling; used in combination with a quantity of solid tobacco, either in a cigarette-like form or in some kind of pod.	Term "heat-not-burn" is not advised for use. If using the term HTP, you may want to give at least one example product.
Practice of using a vaping device	**Vaping** **Nicotine vaping**	Use of a [personal] vaporizer or vaping device. Use of a vaping device containing nicotine.	Does not apply to use of HTPs. "Vaping" can be used for nicotine vaping if it is previously clear that only nicotine vaping is being considered. For non-nicotine vaping, the qualifier "non-nicotine" should always be used. "Vaping" alone can also be used when it is made clear the reference is to either vaping with nicotine or more generally vaping irrespective of specific content.
User of a vaping device	**Vaper,** **Nicotine vaper** ***People who vape/People who use e-cigarettes**	Someone who uses a vaping device, with or without nicotine. Someone who uses a vaping device with nicotine.	"Vaper" can be used for nicotine vaper if it is previously clear that only nicotine vaping is being considered. "Vaper" alone can also be used when it is made clear the reference is to vaping with or without nicotine or the context clearly implies it is with nicotine.
The substance produced by vaping	**Vapour**	The substance produced by vaping.	On first introduction of the term it may be better to say "aerosol, commonly referred to as vapour". Use only when referring to NVPs, not when referring to HTPs. We don't yet have consensus on a term for the emissions of HTPs, but would like to avoid "vapour" if possible.
The liquid used in vaping	**[Nicotine] vaping liquid**	The liquid used for vaping.	Commonly called e-liquid, but not a particularly useful term. Recommend we avoid the term "e-liquid". Can be qualified by core constituent; e.g. "nicotine vaping liquid" or "nicotine liquid" where it can be inferred readily that the liquid is for vaping.
Therapeutic nicotine	**Nicotine replacement therapy (NRT)**	An approved nicotine therapy.	Do not describe vaping as NRT unless it becomes an approved medicine (and only for the forms that are approved medicines).

Concept (products)	Main term	Definitions	Comments
All potentially harm reduced products	**Alternative nicotine product (ANP)**	All potentially harm reduced nicotine products when in a context of possible use outside approved therapeutic use; includes all vaping and any recreational or long-term non-therapeutic use for NRT; includes recreational oral nicotine products like Zyn; may include clean smokeless tobacco.	Where important, make it clear if recreational use of NRT is included. Or specifically excluded. "Clean smokeless tobacco" if included should be defined through examples.
All combusted tobacco products	**Smoked tobacco**	Products in which tobacco is burned to produce smoke for inhaling; includes cigarettes, cigars, bidi, pipes, and some shisha devices.	Some shisha devices do not burn the tobacco (also ANPs like Eclipse), but use a carbon coal to heat it to release the desired nicotine and flavours. If the intention in using the term "smoked tobacco" is to include all shisha, this should be stated. For now, do not take this term to include HTPs. There is a discussion starting on whether certain products such as some HTPs should be included as combusted (as there is some combustion), or heated (as the intent is not to combust the tobacco).
Smokeless tobacco used orally	**Oral tobacco**	Tobacco products in which tobacco leaf or powder is put into the mouth; includes snus, oral snuff, chewing tobacco.	Recommend use "oral" instead of "smokeless" to avoid confusing chew and oral snuff with HTPs. Any use to include nasal snuff should mention nasal snuff explicitly.
Product placed between the upper lip and gum for sub-labial administration	**Tobacco-Free Oral Nicotine Pouches**	Tobacco-free nicotine product in which the pouch containing nicotine is put between the upper lip and gum where nicotine is released.	
	Snus	Powdered tobacco product placed between the lips and gums.	
Combined market for combusted tobacco and other nicotine delivery products	**Nicotine market**	The consumer market for all nicotine delivery products, including cigarettes, other combusted tobacco products, smokeless tobacco, NVPs, HTPs and NRT.	Could be qualified with "recreational" or "therapeutic" where a distinction is important.
Exposure of others	**Second-hand vapour**	Effluent exhaled or emitted during vaping.	"Second-hand aerosol" would seem more correct than "second-hand vapour" because generally particles are present as well as the gaseous component. However, it

Concept (products)	Main term	Definitions	Comments
			may be reasonable to use "second-hand vapour" with suitable qualification.
	Passive vaping	Inhaling second-hand vapour.	Either term "second-hand vapour" or "passive vaping" can be used depending on whether focus is on the effluent or on its inhalation.
Cigarette smoker (definitions consistent with terminology in 4CV1/2 questionnaires)	**Current smoker** ***People who currently smoke**	Someone who has smoked at least 100 cigarettes in lifetime, and currently smokes at least monthly. Corresponds to FR225 = 1-2 or FR225 = 3 and BI345 = 1. (Table D.2)	This definition of current smoker is in line with the eligibility criterion for being recruited as a smoker in the ITC 4C surveys, and seems to be preferred as the meaning of "current smoker" by several of the commenters. In 4CV1, corresponds to FR225 = 1-2 or FR225 = 3 and BI345 = 1. (100 cigarettes was confirmed only if the respondent smoked monthly or less frequently.) The eligibility criterion for 4CV1 was to be smoking daily, weekly, monthly or "at least occasionally", corresponding to FR309v = 1-3. The word "occasional" need only be used when describing the eligibility criteria. The terms "occasional use" or "occasional user "are not advised for use.
	Daily smoker ***People who smoke daily**	A current smoker who currently smokes daily. Corresponds to FR309v = 1 (Table D.2)	
	Non-daily smoker ***People who smoke less than daily**	A current smoker who does not smoke daily.	
	Weekly smoker ***People who smoke weekly**	A current smoker who smokes at least weekly but not daily. Corresponds to FR309v = 2 (Table D.2)	
	Monthly smoker ***People who smoke monthly**	A current smoker who smokes at least monthly but not weekly. Corresponds to FR225 = 3 and BI345 = 1 (Table D.2)	
Users of both cigarettes and NVPs	**Dual user** ***People who vape and smoke**	Someone who both vapes nicotine and smokes cigarettes at least monthly	This term is introduced to remove confusion created by reserving "dual user" to describe someone with daily use of both NVPs and cigarettes. See Borland dual use paper for explanation. You will need to define it when first used as it is not currently understood. Dual use also only refers to using two sources of nicotine.
	Dual daily user	Someone who smokes daily AND uses NVPs daily.	

Concept (products)	Main term	Definitions	Comments
	***People who vape and smoke daily**		Not included in these definitions are respondents who reported using either product less than monthly. They are not treated as current users Also not included are respondents who use vaping devices without nicotine. Groups like these may be given names (which should be defined carefully) specific to their contexts. "Predominant use" will need to be defined before using it. NB: Dual use of smoking and HTP and HTP and VPs will need to be dealt with at some point.
	Dual non-daily user ***People who vape and smoke less than daily**	Someone who smokes weekly or monthly AND uses NVPs weekly or monthly but not daily.	
	Predominant smoker ***People who predominantly smoke**	Someone who smokes daily but uses NVPs weekly or monthly but not daily.	
	Predominant vaper ***People who predominantly vape**	Someone who vapes nicotine daily but smokes weekly or monthly but not daily.	
User of a vaping device	**Vaper** ***People who vape**	Someone who uses a vaping device, with or without nicotine.	Vaping at least weekly was the eligibility criterion to be recruited in 4CV1-4 as a vaper. Qualify by preceding "vaper" with "nicotine" if referring to vaping nicotine and this is not clear. If using the term "current vaper" please use it for someone who vapes nicotine at least monthly and define it as such. At some point we might like to redefine "experimental" to also include past users who are open to trying again. "Ever vaper" will generally be used as an inclusion criteria, or as "only ever tried" to distinguish those who have only ever tried "once or twice" from experimental vapers.
	Daily [nicotine] vaper ***People who vape daily**	A vaper who uses an NVP daily.	
	Weekly [nicotine] vaper ***People who vape weekly**	A vaper who uses an NVP at least weekly, but not daily.	
	Non-daily [nicotine] vaper	A vaper who uses an NVP less often than daily, but at least monthly.	

Concept (products)	Main term	Definitions	Comments
	***People who vape less than weekly**		
	Experimental vaper ***People who vape experimentally**	Someone who uses an NVP less often than monthly.	
	Ever vaper ***People who ever vaped**	Someone who has tried vaping at least once.	
User of a HTP	**HTP user** ***People who use HTPs**	Someone who uses a HTP or HTPs.	Do not use "vaper" for user of HTP. Please only use "user of HTP" or in full without abbreviation if you need to mention users of HTPs. Or "HTP user".
Former smoker	**Recent former smoker** ***People who recently quit** **Quit within the previous two years** ***People who quit within the previous two years**	Someone who has smoked at least 100 cigarettes in lifetime, and is no longer smoking cigarettes. The state of being a former smoker who at the time of the interview had quit within the previous two years.	Qualification with recent and long-term as yet unquantified, but the advice is not to use "recent" for anything more than 6 months or "long term" for anything less than 1 year if possible. Having quit within the previous two years is an eligibility criterion for our study, so we will often be dividing smokes as quit for > or < 2 years.
Former vaper	**Former [nicotine] vaper** ***People who used to vaped**	Someone who vaped at least weekly in the past but currently does not vape.	Deliberate use of different term than for smoking, as all past use is much shorter than for smoking and few have quit vaping after developing a vaper lifestyle. We may eventually want to differentiate former daily from former non-daily vapers. Try not to use "former NVP user". Should be restricted to nicotine vaping, unless it is made clear it is either or non-nicotine.
	Recent former [nicotine] vaper	Someone who used NVPs in the past but has stopped doing so within the previous 2 years (or 24 months).	

Concept (products)	Main term	Definitions	Comments
	***People who used to vape**		
Daily nicotine user	**Daily nicotine user** ***People who use nicotine daily**	Someone who uses a nicotine-containing product daily.	Whether or not to include NRT as a nicotine-containing product could depend on the context.

Table D.2: Cigarette screening variables* FR225, BI345, QA439, and FR309v from the ITC 4CV5 Survey.

*As referred to in Table D.1 to define various types of "cigarette smokers".

Varname	4CV5 Survey Questions
FR225	**Ask all.** How often, if at all, do you CURRENTLY smoke ordinary cigarettes (either factory-made/packet or roll-your-own)? 1 Daily 2 Less than daily, but at least once a week 3 Less than weekly, but at least once a month 4 Less than monthly, but occasionally 5 I have quit smoking 6 I have never been a smoker 8 Refused 9 Don't know *This is an essential question that will help to determine your eligibility. Please do your best to answer. If you can't or don't wish to answer this question, you will not be able to continue.* **If response=4, go to FR142.** **If response=1, 2, 3, 5 or 6, go to FR142v.** **If response=8 or 9, go to BI473, then BI904. (DC: if C, then C-B11.5; if P, then P-C11.5)**
BI345 [A]	**Ask if [P and (FR225v=3, 4 or 5)] OR [C and (FR225v=3, 4 or 5) and (BI345v@LSD=2)].** Have you smoked 100 or more cigarettes over your lifetime? 1 Yes 2 No 8 Refused 9 Don't know *100 cigarettes= 5 packs of 20 cigarettes OR 4 packs of 25 cigarettes.* *This is an essential question that will help to determine your eligibility. Please do your best to answer. If you can't or don't wish to answer this question, you will not be able to continue.* **If response=8 or 9, go to BI473, then BI904. (DC: if C, then C-B11.5; if P, then P-C11.5)**
BI345v	(Derived value – smoked 100 or more cigarettes in lifetime.) If BI345=1, then BI345v=1. If BI345=2, then BI345v=2. If FR225v=1-2, then BI345v=1. If BI345v@LSD=1, then BI345v=1. If FR225v=6, then BI345v=2. 1 Yes 2 No
QA439	**Ask if [(P and (FR304=1 or FR225v=5)) OR (C and (FR304=1 or FR225v=5 or QA342=1-2))] and BI345v=1.** **QA342=2:** How long ago did your CURRENT quit attempt start? **FR304=1 or 2:** How long ago did you reduce your smoking to less than monthly? If you don't know exactly when this happened, give us your best estimate. **Otherwise:** How long ago did you quit smoking?

	01 Less than 1 week ago 02 1-2 weeks ago 03 3-4 weeks ago 04 1-3 months ago 05 4-6 months ago 06 7-12 months ago 07 13-18 months ago 08 19-24 months ago 09 2-3 years ago 10 3-5 years ago 11 More than 5 years ago 88 Refused 99 Don't know *This is an essential question that will help to determine your eligibility. Please do your best to answer. If you can't or don't wish to answer this question, you will not be able to continue.* **If response=88 or 99, go to BI473, then BI904. (DC: if C, then C-B11.7; if P, then P-C11.7)**
FR309v	**Two-digit codes for FR309v were adopted after W2 fieldwork and retroactively implemented for W1 and W2. These allow splitting of the previous category 'Current Occasional/< Weekly Smoker', coded 3, into new categories 'Monthly' and 'Less-than-monthly' smokers, now coded 31 and 32 respectively. Other codes simply have a zero added to the single digit:** *Old -> New* *1 -> 10* *2 -> 20* *3 -> 31 + 32* *4 -> 40* *8 -> 80* *9 -> 90* (Derived variable for all respondents -- cigarette smoking status at current wave, to be used in question filters throughout survey.) 10 Current Daily Smoker (FR225v=1) 20 Current Weekly Smoker (FR225v=2) 31 Current Monthly Smoker (FR225v=3 AND BI345v=1) 32 Current Less-than-monthly Smoker (FR225v=4 AND BI345v=1) 40 Recent Quitter: Identifies as quit in last 24M AND has smoked 100+ lifetime cigs [(FR225v=5 AND BI345v=1 AND QA439=1-8) or (QA342=1-2 and QA439=1-8)] 80 Long-term Quitter: Identifies as quit more than 24M ago AND has smoked 100+ lifetime cigs [(FR225v=5 AND BI345v=1 AND QA439=9-11) or (QA342=1-2 and QA439=9-11)] 90 Non-Smoker: Identifies, or is derived as, never having been a smoker OR has not smoked 100+ lifetime cigs. (FR225v=6 or BI345v=2)

APPENDIX E: Summary of recruitment materials using ITC phrasing vs. standard panel templates

Country	Firm and sample type	Used ITC phasing, image, logo	Used standard panel template	Notes
Pre-Field Interwave Reminder (Recontact only)				
AU	RM (ITC-owned, panel-owned)	√		
	UQ Dedicated Vapers (ITC-owned)	√		
CA	SCSRU (ITC-owned)	√		
	Leger (panel-owned)	**n/a**		Not sent
EN	SCSRU (ITC-owned)	√		
	Ipsos UK (panel-owned)	Partial		Partial = Some ITC phrasing, but no pictures.
	Rakuten (panel-owned)	Partial		
	Rakuten Partners (panel-owned)	Partial		
US	SCSRU (ITC-owned)	√		
	Ipsos UK, Partner(s) (panel-owned)	Partial		
	Ipsos PA (panel-owned)	N/A		
Invitation Email				
AU	RM (ITC-owned, panel-owned) – Recontact, Replenishment	√		
	UQ Dedicated Vapers (ITC-owned) - Recontact only	√		
CA	SCSRU (ITC-owned) – Recontact only	√		
	Leger (panel-owned) – Recontact, Replenishment		√	
EN	SCSRU (ITC-owned) – Recontact only	√		
	Ipsos UK (panel-owned) – Recontact only		√	
	Rakuten (panel-owned) – Recontact, Replenishment	√		
	Rakuten Partners (panel-owned) – Recontact, Replenishment	Partial		
US	SCSRU (ITC-owned) - Recontact only	√		
	Ipsos UK, Partner(s) (panel-owned) – Recontact only		√	
	Ipsos PA (panel-owned) – Recontact and Replenishment	Partial		
	Ipsos PA Partners (panel-owned) – Replenishment only	Partial		
Email Reminder				
AU	RM (ITC-owned, panel-owned) – Recontact, Replenishment	√		
	UQ Dedicated Vapers (ITC-owned) - Recontact only	√		
CA	SCSRU (ITC-owned) – Recontact only	√		
	Leger (panel-owned) – Recontact, Replenishment		√	
EN	SCSRU (ITC-owned) – Recontact only	√		
	Ipsos UK (panel-owned) – Recontact only		√	
	Rakuten (panel-owned) – Recontact, Replenishment	√		
	Rakuten Partners (panel-owned) – Recontact, Replenishment	Partial	√	
US	SCSRU (ITC-owned) - Recontact only	√		
	Ipsos UK, Partner(s) (panel-owned) – Recontact only		√	
	Ipsos PA (panel-owned) – Recontact , Replenishment	√	√	
	Ipsos PA Partners (panel-owned) – Replenishment only		√	

APPENDIX F: Sampling weights, design, and benchmarks

Sampling Weights of the International Tobacco Control (ITC) 4CV Survey

5th Edition (Waves 1–5)

C. Boudreau [†,1,2], Y. Li[2,3], and M.E. Thompson[1,2]

This document describes the various cross-sectional and longitudinal weights for wave 1 (section 1), wave 2 (section 2), wave 3 (section 3), wave 4 (section 4) and wave 5 (section 5) of the ITC Four Country Smoking and Vaping Survey (4CV) Survey. It also provides some guidance on which set of weights should be used depending on the analysis being performed, as well as cautionary notes when analyzing 4CV data (section 6). Lastly, a short description of the various weight constructed in this report is provided on page 78. This description aims to help authors provide a summary of the weight calculation process of the 4CV.

All sampling weights adjust for the oversampling of 18–24 years old tobacco users (at waves 1–3 in US, Canada & England), the oversampling of nicotine vaping product (NVP) users, e-cigarette users and (starting at wave 3) heated tobacco product (HTP) users, sample mis-representation, non-response and other biases. It is thus essential to use weighted data, when preforming any analyses using 4CV data.

Contents

[†]Correspondence: cboudreau@uwaterloo.ca
[1]Dept. of Statistics & Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada
[2]Data Management Core (DMC) – ITC Project, University of Waterloo
[3]Dept. of Psychology, University of Waterloo
[4]This document was created using LaTeX, and last updated on Nov 11, 2025

List of Tables

1 Wave 1 weights

Eighteen sets of cross-sectional weights and 4 sets of longitudinal weights were computed at wave 1 of the 4CV Survey. These cross-sectional weights are listed in table 1, and their computation and how/when they should be used are detailed in section 1.1. Likewise, the longitudinal weights are listed in table 2, and their computation and use are detailed in section 1.2.

All sampling weights for the 4CV Survey were computed using the statistical software R (r-project.org). As mentioned at the beginning of this document, these weights adjust for: 1) oversampling of 18–24 years old tobacco users, 2) the oversampling of nicotine vaping product (NVP) users and e-cigarette users, 3) sample mis-representation, 4) non-response and 5) other biases. **It is thus essential to use weighted data, when performing any analyses using 4CV data**.

Weights	Variable names		
	Main sample⋆	reduced US sample†	AU NVP users‡
Wave 1 cross-sectional inflation weights	kWTS100v	kWTS102v	n/a
Wave 1 cross-sectional inflation weights for NVP users	n/a	n/a	kWTS304v
Rescaled wave 1 cross-sectional weights for cigarette smokers	kWTS201v	kWTS203v	n/a
Rescaled wave 1 cross-sectional weights for NVP users	kWTS301v	kWTS303v	kWTS305v
Rescaled wave 1 cross-sectional weights for dual users	kWTS401v	kWTS403v	kWTS405v
Rescaled wave 1 cross-sectional weights for quitters	kWTS501v	kWTS503v	n/a
Rescaled wave 1 cross-sectional weights for quitters who use NVP	n/a	n/a	kWTS505v
Rescaled wave 1 cross-sectional weights for all tobacco users	kWTS601v	kWTS603v	n/a
Rescaled wave 1 cross-sectional weights for all respondents	kWTS101v	kWTS103v	n/a

⋆ The main sample consists of 12294 respondents (2733 from the US, 3733 from Canada, 4324 from England and 1504 from Australia). It includes all respondents except the 581 Australian respondents from the dedicated CCV sample, the 224 long-term quitters (79 from the US, 95 from Canada, 50 from England, and 0 from Australia) mentioned in the remark section of page 1 and the 87 respondents (26 from the US, 13 from Canada, 48 from England, and 0 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).
† The reduced sample is the main US sample of 2733 respondents minus the 494 US NVP users recruited by Ipsos; hence, the reduced US sample consists of 2239 respondents. Those NVP users were recruited via a non-probability based panel, and those sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired for their specific analyses.
‡ This sample contains all 816 Australian respondents that were using NVP at the time of wave 1. As detailed in section 1.1.4, this was done to allow for weighted analyses including the 581 respondents recruited via the dedicated CCV sample.

Table 1: Cross-sectional sampling weights for wave 1 of the 4CV Survey

Weights	Variable names
Rescaled waves 4C8–4CV1 longitudinal weights	`kWTS967v`
Rescaled waves 4C8.5–4CV1 longitudinal weights (AU only)	`kWTS969v`
Rescaled waves 4C9–4CV1 longitudinal weights	`kWTS971v`
Rescaled waves 4C10–4CV1 longitudinal weights (EN & AU only)	`kWTS973v`

Table 2: Longitudinal sampling weights for wave 1 of the 4CV Survey

1.1 Cross-sectional sampling weights

The 18 sets of cross-sectional weights can be divided into 3 categories:

1. The first 7 sets of cross-sectional weights (see section 1.1.2 and column 2 of table 1) were computed for the main sample. That sample consists of 12294 respondents (2733 from the US, 3733 from Canada, 4324 from England and 1504 from Australia; see table 3). It includes all respondents except the 581 Australian respondents from the dedicated Cancer Council Victoria (CCV) sample[1], the 224 long-term quitters (79 from the US, 95 from Canada, 50 from England, and 0 from Australia) mentioned in section 6 and the 87 respondents (26 from the US, 13 from Canada, 48 from England, and 0 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).

2. The next 7 sets of cross-sectional weights (see section 1.1.3 and column 3 of table 1) were computed for the reduced US sample. This sample is the main US sample of 2733 respondents (mentioned above) minus the 494 NVP users recruited via Ipsos' non-probability Web panel; hence, the reduced US sample consists of 2239 respondents. Those 494 respondents are the only US respondents not recruited using a probability based method, and those sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired. Since this only applies to US respondents, no such weights were computed for respondents from Canada, England or Australia.

3. The last 4 sets of cross-sectional weights (see section 1.1.4 and column 4 of table 1) were computed for Australian respondents using NVP. This sample contains all 816 Australian respondents that were using NVP at the time of wave 1. This was done to allow for weighted analyses including the 581 respondents recruited via the dedicated CCV sample. This concerns Australian respondents only, and no such weights were computed for respondents from the US, Canada or England.

1.1.1 User groups

For all sets of cross-sectional weights, respondents were first divided into 4 user groups (variable `kUserGroup`[2] in the dataset): *i*) cigarette only smokers, *ii*) dual users, *iii*) exclusive nicotine vaping product (NVP) users, and *iv*) quitters (i.e., quit smoking cigarettes within the last 5 years for US,

[1] The Cancer Council Victoria (CCV) sample is a dedicated supplementary sample of Australian respondents who vape at least monthly. These respondents were recruited via referral sampling from online vaper forums and vape stores.

[2] Where 1 = cigarette only, 2 = dual users, 3 = exclusive NVP users, and 4 = quitters; see appendix A.1.

Canada and England, and within the last 2 years for Australia). The numbers of respondents in each of those user groups are given in table 3, and a detailed description of these groups is provided below. In addition to those 4 groups, quitters (i.e., group *iv*) were further divided into 4 sub-groups (variable `kQuitGroup`[3] in the dataset):

iv.a) those who had quit within the last year, but were using NVP (this includes e-cigarettes) at least monthly at the time of data collection

iv.b) those who had quit 1–5 years ago (1–2 years in Australia), but were using NVP (this includes e-cigarettes) at least monthly at the time of data collection

iv.c) those who had quit within the last year, and were using NVP less than monthly or not at all at the time of data collection

iv.d) those who had quit 1–5 years ago (1–2 years in Australia), and were using NVP less than monthly or not at all at the time of data collection.

The following bullet points detail how respondents were classified into the 4 user groups:

- To be classified as a cigarette only smoker (group *i*), a respondent had to (at the time of data collection) smoke cigarettes (at least occasionally), smoked at least 100 cigarettes in their lifetime, and use NVP (this includes e-cigarettes) less than monthly or not at all.

  It should be noted that including occasional smokers, as opposed to at least monthly smokers, is not ideal. This results in a disconnect between our 4CV sample and the benchmark/calibration figures from the various national health surveys used to compute the sampling weights (see variable `kWTS100v` in section 1.1.2). Those benchmark figures typically represent the estimated numbers of at least monthly smokers (who have smoked at least 100 cigarettes in their lifetime). Nevertheless, including those less than monthly but occasional smokers was necessary, as our sample contains 716 such individuals (94 from the US, 322 from Canada, 274 from England and 26 from Australia).

- To be classified as a dual user (group *ii*), a respondent had to (at the time of data collection) smoke cigarettes (at least occasionally), smoked at least 100 cigarettes in their lifetime, and use NVP (this includes e-cigarettes) at least monthly.

  It should be noted that the use of the at "least monthly" criteria for the use of NVP is not ideal, as respondents had to be at least weekly NVP users to be recruited as part of the dedicated subsample of NVP users. This is a consequence that the the various national health surveys, used to compute the sampling weights (see variable `kWTS100v` in section 1.1.2), only allowed to obtain benchmark/calibration figures for monthly usage of NVP. Nevertheless, our sample includes many less than weekly but at least monthly NVP users, as weekly use of NVP was not a criteria to be recruited as part of the main sample of cigarette smokers. Regardless, this is a limitation of the 4CV Survey.

  In addition, the above comment (see definition for group *i*) about including those less than monthly but occasional smokers also applies here.

- To be classified as an exclusive nicotine vaping product (NVP) user (group *iii*), a respondent had to (at the time of data collection) use NVP (this includes e-cigarettes) at least monthly, and satisfy one or more of the following:

[3] Where 1 = quit within the last year and uses NVP, 2 = quit more than 1 year ago and uses NVP, 3 = quit within the last year and does not use NVP, and 4 = quit more than 1 year ago and does not use NVP; see appendix A.1.

- not currently smoking cigarettes
- smoked less than 100 cigarettes in their lifetime
- quit smoking cigarettes more than 5 years ago for the US, Canada and England, or more than the 2 years for Australia

The above comment (see definition for group *ii*) about the use of the at "least monthly" criteria for the use of NVP being a limitation of our study also applies here.

- To be classified as a quitter (group *iv*), a respondent had to (at the time of data collection) have quit smoking cigarettes within the last 5 years for the US, Canada and England, and within the last 2 years for Australia, and smoked at least 100 cigarettes in their lifetime prior to that. Those who were using NVP (this includes e-cigarettes) at least monthly were then further sub-classified in either group *iv.a* or *iv.b* depending on how long ago they had quit smoking; whereas those who were using NVP less than monthly or not at all were then further sub-classified in either group *iv.c* or *iv.d* depending on how long ago they had quit smoking.

  The above comment (see definition for group *ii*) about the use of the at "least monthly" criteria for the use of NVP being a limitation of our study also applies here.

User group†	US		Canada	England	Australia	
	Main‡	Reduced‡			Main‡	NVP‡
Cigarette smokers						
Cigarette only	1347	1347	2179	2664	1147	0
Dual users	980	542	1036	1222	192	292
Total	2327	1889	3215	3886	1339	292
Exclusive NVP/HTP users	37	20	71	17	5	29
Quitters						
NVP/HTP users	126	87	103	181	38	495
Non users	243	243	344	240	122	0
Total	369	330	447	421	160	495
Total	2733	2239	3733	4324	1504	816

† Variables `kUserGroup` and `kQuitGroup` in the dataset
‡ See notes below table 1 and beginning of section 1.1

Table 3: Wave 1 respondents by country and user group

1.1.2 Cross-sectional weights for the main sample

1- Variable `kWTS100v` contains the wave 1 cross-sectional inflation weights for the main sample of 12294 respondents (2733 from the US, 3733 from Canada, 4324 from England and 1504 from Australia; see table 3).

For user groups *i*, *ii* and *iv*, respondents were further subdivided based on gender, age, geographic region, ethnicity (US only), education (except for Canada) and language (Canada only). As can be seen in table 1, the number of exclusive NVP users (group *iii*) is much smaller than that of the other

user groups. It was thus not possible to create as many subgroups. In Canada and the US, exclusive NVP users were further subdivided based on gender and age (i.e., 18–24 vs. > 24). For Australia and England, exclusive NVP users were not further subdivided. This yielded the following cross-tabs: user group × gender, user group × age group, user group × geographic region, user group × ethnicity (US only), user group × education (except for Canada) and user group × language (Canada only); where some of those subgroups/cells (in particular exclusive NVP users) were collapsed because they contained too few respondents.

Data from various national health surveys were then used to obtain benchmark/calibration figures (e.g., estimated number of individuals that are dual users) for each of the above mentioned cross-tabs. In the US, the 2016 National Health Interview Survey (NHIS) was used; whereas the 2015 Canadian Tobacco Alcohol and Drugs Survey (CTADS) and the 2016 National Drug Strategy Household Survey (NDSHS) were used for Canada and Australia. In England, the 2015 Opinions and Lifestyle Survey (OLS) was combined with waves 117–122 (Jun–Nov 2016) of the Smoking Toolkit Study (STS) to obtain the calibration figures. These figures are given in appendix A.2. A raking procedure (see appendix A.3) was then applied to calibrate the weights using the above mentioned cross-tabs; this was done separately for each country.

These weights are designed to make respondents in each of the four groups representative of the corresponding population at the time of wave 1 data collection. For example, the `kWTS100v` weights of the 1036 Canadian dual users are designed to make them representative of the Canadian population of dual users at the time of data collection; likewise for the other countries and the other groups. If interests lie in a target population that consists of two or more of the four user groups, the `kWTS100v` weights are still appropriate. For example, when studying Canadian cigarette smokers, one can simply combine the `kWTS100v` weights of the 2179 cigarette only users with those of the 1036 dual users (for a total of 3215 respondents in the analysis), and assign a weight of 0 to respondents in the other two user groups.

Last but not least, since these are inflation/un-rescaled weights, they should not be used in analyses involving two or more countries. The various rescaled weights (i.e., variables `kWTS101v` to `kWTS601v`) described below were created especially for such multi-country analyses; see section 6.2 for more information on inflation versus rescaled weights.

2- Variable `kWTS201v` contains the rescaled wave 1 cross-sectional weights for the 10767 (2327 from the US, 3215 from Canada, 3886 from England and 1339 from Australia; see table 3) respondents who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS100v`) of those 10767 respondents rescaled to sum to sample size in each country (i.e., 2327 in the US, 3215 in Canada, 3886 in England and 1339 in Australia). These weights are designed to make these 3215 Canadian cigarette smokers representative of the Canadian population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 1 data collection; likewise for the US, England and Australia.

3- Variable `kWTS301v` contains the rescaled wave 1 cross-sectional weights for the 4008 (1143 from the US, 1210 from Canada, 1420 from England and 235 from Australia; see table 3) respondents who were at least monthly NVP users at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS100v`) of those 4008 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1210 Canadian NVP users representative of the Canadian population of at least monthly NVP users at the time of wave 1 data collection; likewise for the US, England and Australia.

4- Variable `kWTS401v` contains the rescaled wave 1 cross-sectional weights for the 3430 (980 from the US, 1036 from Canada, 1222 from England and 192 from Australia; see table 3) respondents who were dual users at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS100v`) of those 3430 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1036 Canadian dual users representative of the Canadian population of dual users at the time of wave 1 data collection; likewise for the US, England and Australia.

5- Variable `kWTS501v` contains the rescaled wave 1 cross-sectional weights for the 1397 (369 from the US, 447 from Canada, 421 from England and 160 from Australia; see table 3) quitters (within the last 5 years) at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS100v`) of those 1397 respondents rescaled to sum to sample size in each country (i.e., 369 in the US, 447 in Canada, 421 in England and 160 in Australia).

Though these weights are designed to make these 447 Canadian quitters as representative as possible of the Canadian population of quitters at the time of wave 1 data collection, the vast majority of those quitters were initially recruited as smokers. Consequently, they are an imperfect sample when it comes to be representative of the whole population of quitters; see section 6.3 on the representativeness of quitters in the 4CV sample. The same cautionary note also applies to the US, England and Australia. Furthermore, the US, Canadian and English samples consist of those who have quit smoking within the last five years, whereas the Australian sample consists of those who have quit smoking within the last two years. Since the target populations are different, care must be taken when comparing Australian quitters with those of the other three countries.

6- Variable `kWTS601v` contains the rescaled wave 1 cross-sectional weights for the 11345 (2490 from the US, 3389 from Canada, 4084 from England and 1382 from Australia; see table 3) respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users[4]) at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS100v`) of those 11345 respondents rescaled to sum to sample size in each country. These weights are designed to make these 3389 Canadian tobacco users representative of the Canadian population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 1 data collection; likewise for the US, England and Australia.

7- Variable `kWTS101v` contains the rescaled wave 1 cross-sectional weights for the main sample of 12294 respondents (2733 from the US, 3733 from Canada, 4324 from England and 1504 from Australia; see table 3). These are simply the wave 1 cross-sectional inflation weights (variable `kWTS100v`) of those 12294 respondents rescaled to sum to sample size in each country (i.e., 2733 in the US, 3733 in Canada, 4324 in England and 1504 in Australia). These weights are designed to make these 3733 Canadian tobacco users and quitters representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 1 data collection; likewise for the US, England and Australia (within the last 2 years).

In addition to the warning about the representativeness of quitters in the 4CV sample (see variable `kWTS501v` above), it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the `kWTS101v` weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (e.g.,

[4]Note that some of those NVP users are individuals who have quit smoking traditional cigarette.

linear regression and logistic regression) from that same joint population is probably much more questionable.

1.1.3 Cross-sectional weights for the reduced US sample

As mentioned at the beginning of section 1.1, the following 7 sets of weights only apply to the US sample.

1- Variable `kWTS102v` contains the wave 1 cross-sectional inflation weights for the reduced US sample of 2239 respondents. As noted below table 1, the reduced sample is the main sample of 2733 respondents minus the 494 US NVP users recruited by Ipsos (identifiable via the variables `kOWNERID` and `kEC309v`[5] in the dataset). Those at least monthly NVP users were recruited via a non-probability based panel, and the 7 sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired for their specific analyses.

These weights were computed the same way as the wave 1 cross-sectional inflation weights for the main sample (see variable `kWTS100v` in section 1.1.2).

2- Variable `kWTS203v` contains the rescaled wave 1 cross-sectional weights for the 1889 respondents (see table 3) from the reduced US sample who were cigarette smokers at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS102v`) of those 1889 respondents rescaled to sum to sample size (i.e, 1889). These weights are designed to make these 1889 US cigarette smokers representative of the US population of cigarette smokers at the time of wave 1 data collection. In other words, variable `kWTS203v` is the analogue of variable `kWTS201v`, but for the reduced sample.

3- Variable `kWTS303v` is the analogue of variable `kWTS301v`, but for the reduced US sample of 649 respondents who were at least monthly NVP users at the time of wave 1 data collection.

4- Variable `kWTS403v` is the analogue of variable `kWTS401v`, but for the reduced US sample of 542 respondents who were dual users at the time of wave 1 data collection.

5- Variable `kWTS503v` is the analogue of variable `kWTS501v`, but for the reduced US sample of 330 respondents who were quitters at the time of wave 1 data collection. Hence, the cautionary notes listed for variable `kWTS501v` also apply here.

6- Variable `kWTS603v` is the analogue of variable `kWTS601v`, but for the reduced US sample of 1996 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 1 data collection.

7- Variable `kWTS103v` is the analogue of variable `kWTS101v`, but for the reduced US sample of 2239 respondents. Hence, the cautionary notes listed for variable `kWTS101v` also apply here.

1.1.4 Cross-sectional weights for the Australian NVP users

As mentioned at the beginning of section 1.1, the following 4 sets of weights only apply to the Australian sample of 816 respondents (235 from the main sample and 581 from the CCV sample) that were using NVP at least monthly at the time of wave 1. No such weights were computed for respondents from the US, Canada, England or for Australian respondents that are not at least monthly NVP users.

[5] Use !(`kOWNERID`=5 and `kEC309v` $\in \{1, 2\}$) to exclude the 494 US NVP users recruited by Ipsos, where ! is the logic symbol for not.

It should first be mentioned that CCV vaper sample (identifiable via the variable `kOWNERID`[6] in the dataset) consists of self-selected individuals recruited mainly through online sites. They are a group of vaper activists and early adopters. They are thus not a representative sample of vaper users in Australia. This is why those respondents are excluded from the 7 sets of weights computed in section 1.1.2. Nevertheless, respondents from the CCV sample can be of scientific interest, and some data users might want to include them in their analyses. This is why the following 4 sets of weights were computed.

The initial plan was to compute weights for the entire/full sample of 2085 Australian respondents. Though this was technically possible, those weights would have been unstable. It was thus decided to instead compute weights for the 816 respondents that were using NVP at least monthly at the time of wave 1. There are 617 quitters (individuals who have quit cigarette smoking within the last 2 years) in the full sample. Of those, only 122 (or 19.7%) are not (at least monthly) NVP users. However, according to the 2016 NDSHS data, 96.4% of those who quit within the last two years are not using NVP. This huge difference between the 4CV sample of quitters and the 2016 NDSHS target data results in the weights of quitters who do not use NVP being extremely large (the average weight of a quitter who does not use NVP would have been about 200 times that of a quitter who uses NVP); making the weights unstable.

1- Variable `kWTS304v` contains the wave 1 cross-sectional inflation weights for the Australian sample of 816 respondents (235 from the main sample and 581 from the CCV sample) that were using NVP at least monthly at the time of wave 1.

   These weights were computed in a very similar way as the wave 1 cross-sectional inflation weights for the main sample (see variable `kWTS100v` in section 1.1.2). However, respondents were first divided into 3 (dual users, exclusive NVP users and quitters using NVP) instead of 4 user groups. The weights were then calibrated (user group × gender, user group × age group, and the 2 quitter sub-groups) based on figures from the 2016 NDSHS.

2- Variable `kWTS305v` contains the rescaled wave 1 cross-sectional weights for the 816 (235 from the main sample and 581 from the CCV sample; see table 3) Australian respondents who were at least monthly NVP users at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `kWTS304v`) of those 816 respondents rescaled to sum to sample size (i.e, 816). These weights are designed to make these 816 NVP users representative of the Australian population of at least monthly NVP users at the time of wave 1 data collection. In other words, variable `kWTS305v` is the analogue of variable `kWTS301v`, but for all 816 Australian NVP users.

3- Variable `kWTS405v` is the analogue of variable `kWTS401v`, but for the 292 (192 from the main sample and 100 from the CCV sample; see table 3) Australian respondents who were dual users at the time of wave 1 data collection.

4- Variable `kWTS505v` contains the rescaled wave 1 cross-sectional weights for the 495 (38 from the main sample and 457 from the CCV sample; see table 3) Australian respondents who had quit cigarette smoking but were using NVP at least monthly at the time of wave 1 data collection.

1.2 Longitudinal sampling weights

As mentioned at the beginning of section 1, 4 sets of longitudinal weights were computed at wave 1 of the 4CV Survey. They are listed in table 2, and their computation and use are detailed below. It should

[6] Use `kOWNERID`=7 to select the 581 Australian respondents from the dedicated CCV sample.

be noted that all longitudinal weights are designed to make the sample representative at the time of baseline wave, and not at the time of the current wave (i.e., wave 1 in the present case). For example, the waves 4C8–4CV1 longitudinal weights (variable `kWTS967v` below) are designed to make the sample representative of the population of smokers at the time of wave 4C8 data collection. If the aim is to ensure that the sample is representative at the time of wave 1 data collection, then the cross-sectional sampling weights computed in section 1.1 should be used instead.

1- Variable `kWTS967v` contains the waves 4C8–4CV1 (where wave 4C8 denote the 8th wave of the ITC 4C Survey) longitudinal weights for the 1202 respondents (258 from the US, 467 from Canada, 240 from England, and 237 from Australia) who completed the wave 4C8 survey, and were successfully retained and interviewed at wave 1 of the 4CV Survey (and all the waves in between). These weights are designed to make these 467 Canadian smokers (and quitters) representative of the Canadian population of smokers at the time of wave 4C8 data collection; likewise for the US, England and Australia.

2- Variable `kWTS969v` contains the waves 4C8.5–4CV1 longitudinal weights for the 276 Australian respondents who completed the wave 4C8.5 survey, and were successfully retained and interviewed at wave 1 of the 4CV Survey (and all the waves in between). These weights are designed to make these 276 smokers (and quitters) representative of the Australian population of smokers at the time of wave 4C8.5 data collection.

3- Variable `kWTS971v` contains the waves 4C9–4CV1 (where wave 4C9 denote the 9th wave of the ITC 4C Survey) longitudinal weights for the 2662 respondents (1400 from the US, 661 from Canada, 269 from England, and 332 from Australia) who completed the wave 4C9 survey, and were successfully retained and interviewed at wave 1 of the 4CV Survey (and all the waves in between). These weights are designed to make these 1400 US smokers (and quitters) representative of the American population of smokers at the time of wave 4C9 data collection; likewise for the Canada, England and Australia.

4- Variable `kWTS973v` contains the waves 4C10–4CV1 (where wave 4C10 denote the 10th wave of the ITC 4C Survey) longitudinal weights for the 819 respondents (304 from England, and 515 from Australia) who completed the wave 4C10 survey, and were successfully retained and interviewed at wave 1 of the 4CV Survey. These weights are designed to make these 304 English smokers (and quitters) representative of the English population of smokers at the time of wave 4C10 data collection; likewise for Australia.

Note: It was decided to no longer compute waves 4C1–4CV1, 4C2–4CV1, ..., 4C7–4CV1 longitudinal weights. There are essentially 3 reasons for this: 1) before wave 4C8, all interviews were conducted by phone, whereas (practically[7]) all interviews at wave 1 were done online, 2) the tobacco landscape in the 4 countries has changed quite a lot since waves 1–7 of the 4C Survey took place, and 3) given the high level of attrition of cohorts 1–7 respondents between waves 4C8 and 1, the wisdom of conducting longitudinal analyses of those cohorts is becoming more and more questionable.

2 Wave 2 weights

Eighteen sets of cross-sectional weights and 17 sets of longitudinal weights were computed at wave 2 of the 4CV Survey. These cross-sectional weights are listed in table 4, and their computation and

[7] 46 US cohort respondents where exceptionally allowed to complete the wave 1 survey by phone.

how/when they should be used are detailed in section 2.1. Likewise, the longitudinal weights are listed in table 5, and their computation and use are detailed in section 2.2.

All sampling weights for the 4CV Survey were computed using the statistical software R (r-project.org). As mentioned at the beginning of this document, these weights adjust for: 1) oversampling of 18–24 years old tobacco users, 2) the oversampling of nicotine vaping product (NVP) users and e-cigarette users, 3) sample mis-representation, 4) non-response and 5) other biases. **It is thus essential to use weighted data, when performing any analyses using 4CV data**.

Weight	Variable Names		
	Main sample⋆	reduced US sample†	AU NVP users‡
Wave 2 cross-sectional inflation weights	lWTS100v	lWTS102v	n/a
Wave 2 cross-sectional inflation weights for NVP users	n/a	n/a	lWTS304v
Rescaled wave 2 cross-sectional weights for cigarette smokers	lWTS201v	lWTS203v	n/a
Rescaled wave 2 cross-sectional weights for NVP users	lWTS301v	lWTS303v	lWTS305v
Rescaled wave 2 cross-sectional weights for dual users	lWTS401v	lWTS403v	lWTS405v
Rescaled wave 2 cross-sectional weights for quitters	lWTS501v	lWTS503v	n/a
Rescaled wave 2 cross-sectional weights for quitters who use NVP	n/a	n/a	lWTS505v
Rescaled wave 2 cross-sectional weights for all tobacco users	lWTS601v	lWTS603v	n/a
Rescaled wave 2 cross-sectional weights for all respondents	lWTS101v	lWTS103v	n/a

⋆ The main sample consists of 12987 respondents (2846 from the US, 3778 from Canada, 4848 from England and 1515 from Australia). It includes all respondents except the 625 Australian respondents from the dedicated CCV sample, the 7 long-term quitters (2 from the US and 5 from Canada) who are not NVP users and the 148 respondents (36 from the US, 51 from Canada, 53 from England, and 8 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation). Also excluded are an additional 16 respondents from the CCV sample. These respondents are exclusive NVP users recruited at wave 1, and who have quit vaping between waves 1 and 2. Since they no longer meet the eligibility criteria, no cross-sectional weights were computed for them. However, longitudinal weights are designed to make the sample representative at the time of wave 1, and thus longitudinal weights were computed for those 16 respondents.
† The reduced sample is the main US sample of 2846 respondents minus the 129 NVP users recruited by Ipsos at wave 1 and successfully recontacted at wave 2; hence, the reduced US sample consists of 2717 respondents.
‡ This sample contains all 901 Australian respondents that were vaping at the time of wave 2. As detailed in section 2.1.3, this was done to allow for weighted analyses including the 625 respondents recruited via the dedicated CCV sample.

Table 4: Cross-sectional sampling weights for wave 2 of the 4CV Survey

Weight	Variable Names		
Rescaled waves 4C9–2 longitudinal weights	lWTS971v		
Rescaled waves 4C10–2 longitudinal weights (EN and AU only)	lWTS973v		
	Main sample	reduced US sample	AU NVP users
Rescaled waves 1–2 longitudinal weights for cigarette smokers	lWTS221v	lWTS261v	n/a
Rescaled waves 1–2 longitudinal weights for NVP users	lWTS321v	lWTS361v	lWTS381v
Rescaled waves 1–2 longitudinal weights for dual users	lWTS421v	lWTS461v	lWTS481v
Rescaled waves 1–2 longitudinal weights for quitters	lWTS521v	lWTS561v	n/a
Rescaled waves 1–2 longitudinal weights for quitters who use NVP	n/a	n/a	lWTS581v
Rescaled waves 1–2 longitudinal weights for all tobacco users	lWTS621v	lWTS661v	n/a
Rescaled waves 1–2 longitudinal weights for all respondents	lWTS121v	lWTS161v	n/a

Table 5: Longitudinal sampling weights for wave 2 of the 4CV Survey

2.1 Cross-sectional sampling weights

As with wave 1, the 18 sets of cross-sectional weights can be divided into 3 categories:

1. The first 7 sets of cross-sectional weights (see section 2.1.1 and column 2 of table 4) were computed for the main sample. That sample consists of 12987 respondents (2846 from the US, 3778 from Canada, 4848 from England and 1515 from Australia; see table 6). It includes all respondents except the 625 Australian respondents from the dedicated CCV sample, the 7 long-term quitters (2 from the US and 5 from Canada) who are not NVP users and the 148 respondents (36 from the US, 51 from Canada, 53 from England, and 8 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation). Also excluded are an additional 16 respondents from the CCV sample. These respondents are exclusive NVP users recruited at wave 1, and who have quit vaping between waves 1 and 2. Since they no longer meet the eligibility criteria, no cross-sectional weights were computed for them. However, longitudinal weights are designed to make the sample representative at the time of wave 1, and thus longitudinal weights were computed for those 16 respondents.

2. The next 7 sets of cross-sectional weights (see section 2.1.2 and column 3 of table 4) were computed for the reduced US sample. This sample is the main US sample of 2846 respondents (mentioned above) minus the 129 NVP users recruited via Ipsos' non-probability Web panel at wave 1 and successfully recontacted at wave 2; hence, the reduced US sample consists of 2717 respondents. Those 129 respondents are the only US respondents not recruited using a probability based method, and those sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired. Since this only applies to US respondents, no such weights were computed for respondents from Canada, England or Australia.

3. The last 4 sets of cross-sectional weights (see section 2.1.3 and column 4 of table 4) were computed for Australian respondents using NVP at least monthly. This sample contains all 901 Australian respondents that were vaping at the time of wave 2. This was done to allow for weighted analyses including the 625 respondents recruited via the dedicated CCV sample. This concerns Australian respondents only, and no such weights were computed for respondents from the US, Canada or England.

As in wave 1, the first step in the computation of all sets of wave 2 cross-sectional weights consisted in dividing respondents into 4 user groups (variable `lUserGroup` in the dataset). Likewise, quitters were then further divided into 4 sub-groups (variable `lQuitGroup` in the dataset). The numbers of respondents in each of those user groups are given in table 6, and a detailed description of these groups was provided in section 1.1.1.

2.1.1 Cross-sectional weights for the main sample

1- Variable `lWTS100v` contains the wave 2 cross-sectional inflation weights for the main sample of 12987 respondents (2846 from the US, 3778 from Canada, 4848 from England and 1515 from Australia; see table 6). Computation of those weights followed the same steps and used the same raking algorithm (see appendix A.3) as the wave 1 cross-sectional inflation weights for the main sample (variable `kWTS100v` of section 1.1.2). The only differences being:

   i) The wave 2 cross-sectional inflation weights are designed to make the sample representative of the target population at the time of wave 2. Consequently, for cohort respondents, age at wave 2

User group†	US		Canada	England	Australia	
	Main‡	Reduced‡			Main‡	NVP‡
Cigarette smokers						
Cigarette only	1389	1333	2019	2393	1061	0
Dual users	752	709	954	1848	192	248
Total	2141	2042	2973	4241	1253	248
Exclusive NVP/HTP users	111	103	55	30	9	67
Quitters						
NVP/HTP users	229	217	174	336	75	586
Non users	365	355	576	241	178	0
Total	594	572	750	577	253	586
Total	2846	2717	3778	4848	1515	901

† Variables `lUserGroup` and `lQuitGroup` in the dataset
‡ See notes below table 4 and beginning of section 2.1

Table 6: Wave 2 respondents by country and user group.

was used instead of their age when they were first recruited. Likewise, variables `lUserGroup` and `lQuitGroup` were used instead of `kUserGroup` and `kQuitGroup`, and similarly for education and geographic region. It is thus possible for a respondent to transition from being a cigarette only smoker (group *i*) at wave 1 and to being classified as a dual user (group *ii*) at wave 2. Other types of transitions are also possible. Likewise, it is possible for respondents to move between geographic regions or increase his/her education level.

ii) Though the same cross-tabs were used (i.e., user group × gender, user group × age group, user group × geographic region, user group × ethnicity (US only), user group × education (except for Canada) and user group × language (Canada only), the collapsing of cells (in particular exclusive NVP users) because they contained too few respondents was slightly different.

iii) The same 5 national health surveys were used; i.e., the National Health Interview Survey (NHIS) in the US, the Canadian Tobacco Alcohol and Drugs Survey (CTADS) in Canada, the National Drug Strategy Household Survey (NDSHS) in Australia, and the Opinions and Lifestyle Survey (OLS) combined with the Smoking Toolkit Study (STS) in England. However, data from the 2017 NHIS was used instead of the 2016 one; likewise, data from the 2017 CTADS was used instead of the 2015 one, and data from the 2017 OLS (instead of the 2015 OLS) was combined with waves 137—142 (Feb—Jul 2017) of the Smoking Toolkit Study in England. For Australia, the 2016 NDSHS was used at both waves 1 and 2. See the "Wave 2" column of the various tables in appendix A.2.

These weights are designed to make respondents in each of the four groups representative of the corresponding population at the time of wave 2 data collection. For example, the `lWTS100v` weights of the 954 Canadian dual users are designed to make them representative of the Canadian population of dual users at the time of data collection; likewise for the other countries and the other groups. If interests lie in a target population that consists of two or more of the four user groups, the `lWTS100v` weights are still appropriate. For example, when studying Canadian cigarette smokers, one can

simply combine the lWTS100v weights of the 2019 cigarette only users with those of the 954 dual users (for a total of 2973 respondents in the analysis), and assign a weight of 0 to respondents in the other two user groups.

Last but not least, since these are inflation/un-rescaled weights, they should not be used in analyses involving two or more countries. The various rescaled weights (i.e., variables lWTS101v to lWTS601v) described below were created especially for such multi-country analyses; see section 6.2 for more information on inflation versus rescaled weights.

2- Variable lWTS201v contains the rescaled wave 2 cross-sectional weights for the 10608 (2141 from the US, 2973 from Canada, 4241 from England and 1253 from Australia; see table 6) respondents who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable lWTS100v) of those 10608 respondents rescaled to sum to sample size in each country (i.e., 2141 in the US, 2973 in Canada, 4241 in England and 1253 in Australia). These weights are designed to make these 2973 Canadian cigarette smokers representative of the Canadian population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 2 data collection; likewise for the US, England and Australia.

3- Variable lWTS301v contains the rescaled wave 2 cross-sectional weights for the 4765 (1092 from the US, 1183 from Canada, 2214 from England and 276 from Australia; see table 6) respondents who were at least monthly NVP users at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable lWTS100v) of those 4765 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1183 Canadian NVP users representative of the Canadian population of at least monthly NVP users at the time of wave 2 data collection; likewise for the US, England and Australia.

4- Variable lWTS401v contains the rescaled wave 2 cross-sectional weights for the 3746 (752 from the US, 954 from Canada, 1848 from England and 192 from Australia; see table 6) respondents who were dual users at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable lWTS100v) of those 3746 respondents rescaled to sum to sample size in each country. These weights are designed to make these 954 Canadian dual users representative of the Canadian population of dual users at the time of wave 2 data collection; likewise for the US, England and Australia.

5- Variable lWTS501v contains the rescaled wave 2 cross-sectional weights for the 2174 (594 from the US, 750 from Canada, 577 from England and 253 from Australia; see table 6) quitters (within the last 5 years) at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable lWTS100v) of those 2174 respondents rescaled to sum to sample size in each country. The cautionary notes of variable kWTS501v and of section 6.3 still apply here.

6- Variable lWTS601v contains the rescaled wave 2 cross-sectional weights for the 11627 (2481 from the US, 3202 from Canada, 4607 from England and 1337 from Australia; see table 6) respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users[8]) at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable lWTS100v) of those 11627 respondents rescaled to sum to sample size in each country. These weights are designed to make these 3202 Canadian tobacco users representative of the Canadian population

[8]Note that some of those NVP users are individuals who have quit smoking traditional cigarette.

of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection; likewise for the US, England and Australia.

7- Variable `lWTS101v` contains the rescaled wave 2 cross-sectional weights for the main sample of 12987 respondents (2846 from the US, 3778 from Canada, 4848 from England and 1515 from Australia; see table 6). These are simply the wave 2 cross-sectional inflation weights (variable `lWTS100v`) of those 12987 respondents rescaled to sum to sample size in each country (i.e., 2846 in the US, 3778 in Canada, 4848 in England and 1515 in Australia). These weights are designed to make these 3778 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters at the time of wave 2 data collection; likewise for the US, England and Australia (within the last 2 years).

   As mentioned in the description of variable `kWTS101v`, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the `lWTS101v` weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (e.g., linear regression and logistic regression) from that same joint population is probably much more questionable.

2.1.2 Cross-sectional weights for the reduced US sample

As mentioned at the beginning of section 2.1, the following 7 sets of weights only apply to the US sample.

1- Variable `lWTS102v` contains the wave 2 cross-sectional inflation weights for the reduced US sample of 2717 respondents. As noted below table 4, the reduced sample is the main sample of 2846 respondents minus the 129 NVP users recruited by Ipsos at wave 1 (see section 1.1.3) and retained at wave 2. Those NVP users are identifiable via the variable `lOWNERID`[9] in the dataset. Those at least monthly NVP users were recruited via a non-probability based panel, and the 7 sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired for their specific analyses.

   These weights were computed the same way as the wave 2 cross-sectional inflation weights for the main sample (see variable `lWTS100v` in section 2.1.1).

2- Variable `lWTS203v` contains the rescaled wave 2 cross-sectional weights for the 2042 respondents (see table 6) from the reduced US sample who were cigarette smokers at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable `lWTS102v`) of those respondents rescaled to sum to sample size (i.e, 2042). These weights are designed to make these 2042 US cigarette smokers representative of the US population of cigarette smokers at the time of wave 2 data collection. In other words, variable `lWTS203v` is the analogue of variable `lWTS201v`, but for the reduced sample.

3- Variable `lWTS303v` is the analogue of variable `lWTS301v`, but for the reduced US sample of 1029 respondents who were at least monthly NVP users at the time of wave 2 data collection.

4- Variable `lWTS403v` is the analogue of variable `lWTS401v`, but for the reduced US sample of 709 respondents who were dual users at the time of wave 2 data collection.

[9] Use `lOWNERID`!=05 to exclude the 129 US NVP users recruited by Ipsos, where != means not equal to.

5- Variable lWTS503v is the analogue of variable lWTS501v, but for the reduced US sample of 572 respondents who were quitters at the time of wave 2 data collection.

6- Variable lWTS603v is the analogue of variable lWTS601v, but for the reduced US sample of 2362 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection.

7- Variable lWTS103v is the analogue of variable lWTS101v, but for the reduced US sample of 2717 respondents. Hence, the cautionary notes listed for variable lWTS101v also apply here.

2.1.3 Cross-sectional weights for Australian NVP users

As mentioned at the beginning of section 2.1, the following 4 sets of weights only apply to the Australian sample of 901 respondents (276 from the main sample and 625 from the CCV sample; see table 6) that were using NVP (at least monthly) at the time of wave 2. No such weights were computed for respondents from the US, Canada, England or for Australian respondents that are not at least monthly NVP users.

Recall from section 1.1.4 that the CCV vaper sample (identifiable via the variable lOWNERID[10] in the dataset) consists of self-selected individuals recruited mainly through online sites. They are a group of vaper activists and early adopters. They are thus not a representative sample of vaper users in Australia. This is why those respondents are excluded from the 7 sets of weights computed in section 1.1.2. Nevertheless, respondents from the CCV sample can be of scientific interest, and some data users might want to include them in their analyses. This is why the following 4 sets of weights were computed.

1- Variable lWTS304v contains the wave 2 cross-sectional inflation weights for the Australian sample of 901 respondents (276 from the main sample and 625 from the CCV sample; see table 6) that were using NVP at least monthly at the time of wave 2.

These weights were computed in a very similar way as the wave 2 cross-sectional inflation weights for the main sample (see variable lWTS100v in section 2.1.1). However, respondents were first divided into 3 (dual users, exclusive NVP users and quitters using NVP) instead of 4 user groups. The raking algorithm was then used to calibrate the weights on user group × gender, user group × age group, and the 2 quitter sub-groups (i.e., quit within the last year and uses NVP, and quit more than 1 year ago and uses NVP).

2- Variable lWTS305v contains the rescaled wave 2 cross-sectional weights for the 901 Australian respondents (276 from the main sample and 625 from the CCV sample) who were at least monthly NVP users at the time of wave 2 data collection. These are simply the wave 2 cross-sectional inflation weights (variable lWTS304v) of those respondents rescaled to sum to sample size (i.e, 901). These weights are designed to make these 901 NVP users representative of the Australian population of at least monthly NVP users at the time of wave 2 data collection. In other words, variable lWTS305v is the analogue of variable lWTS301v, but for all 901 Australian at least monthly NVP users.

3- Variable lWTS405v is the analogue of variable lWTS401v, but for the 248 (192 from the main sample and 56 from the CCV sample; see table 6) Australian respondents who were dual users at the time of wave 2 data collection.

[10] Use lOWNERID $\in \{07, 70\}$ to select the 625 respondents from the dedicated CCV sample.

4- Variable `lWTS505v` contains the rescaled wave 2 cross-sectional weights for the 586 (75 from the main sample and 511 from the CCV sample; see table 6) Australian respondents who had quit cigarette smoking but were using NVP at least monthly at the time of wave 2 data collection.

2.2 Longitudinal sampling weights

As mentioned at the beginning of section 2, 17 sets of longitudinal weights were computed at wave 2 of the 4CV Survey. They are listed in table 5, and their computation and use are detailed below. It should be noted that all longitudinal weights are designed to make the sample representative at the time of the baseline wave, and not at the time of the current wave (i.e., wave 2 in the present case). For example, the rescaled waves 1–2 longitudinal weights for the main sample (variable `lWTS121v` below) are designed to make the sample representative of the population of tobacco users and quitters representative at the time of wave 1 data collection. If the aim is to ensure that the sample is representative at the time of wave 2 data collection, then the cross-sectional sampling weights computed in section 2.1 should be used instead.

2.2.1 Longitudinal weights for 4C respondents

The following 2 sets of weights were computed for the respondents that were recruited as part of the ITC Four Country (4C) Survey; the predecessor of the current 4CV Survey.

1- Variable `lWTS971v` contains the waves 4C9–1 (where wave 4C9 denote the 9^{th} wave of the 4C Survey) longitudinal weights for the 1749 respondents (906 from the US, 429 from Canada, 180 from England, and 234 from Australia) who completed the wave 4C9 survey, and were successfully retained and interviewed at wave 2 (and all the waves in between). These weights are designed to make these 906 US smokers (and quitters) representative of the American population of smokers at the time of wave 4C9 data collection; likewise for the Canada, England and Australia.

2- Variable `lWTS973v` contains the waves 4C10–1 (where wave 4C10 denote the 10^{th} wave of the 4C Survey) longitudinal weights for the 557 respondents (206 from England, and 351 from Australia) who completed the wave 4C10 survey, and were successfully retained and interviewed at wave 2. These weights are designed to make these 206 English smokers (and quitters) representative of the English population of smokers at the time of wave 4C10 data collection; likewise for Australia.

Note that 4C8–1 and 4C8.5–1 weights were computed at wave 1 (see table 2 and section 1.2). No such weights were computed at wave 2, as too few responds who complete the waves 8 and/or 8.5 surveys were retained at wave 2.

2.2.2 Waves 1–2 longitudinal weights for the main sample

Waves 1–2 longitudinal inflation weights were computed for the 5632 respondents (1198 from the US, 1865 from Canada, 1719 from England and 850 from Australia) from the wave 1 main sample that were successfully recontacted at wave 2. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

User group†	US		Canada	England	Australia	
	Main	Reduced			Main	NVP
Cigarette only	708	708	1113	1129	653	0
Dual users	246	198	487	396	95	135
Exclusive NVP users	17	0	30	0	0	16
Quitters						
NVP users	74	58	55	85	29	304
Non users	153	153	180	109	73	0
Total	227	211	235	194	102	304
Total	1198	1117	1865	1719	850	455

† Variables `kUserGroup` and `kQuitGroup` in the dataset

Table 7: Wave 1 respondents successfully recontacted at wave 2 by user group.

These waves 1–2 longitudinal inflation weights are the wave 1 cross-sectional inflation weights (i.e., variable `kWTS100v`, computed in section 1.1.2) adjusted for attrition between waves 1 and 2. Consequently, computation of those longitudinal inflation weights followed the same steps as described in section 1.1.2. Since these weights are meant to make these 5632 respondents representative of the population at the time of wave 1 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 1 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 2 cross-sectional weights for the main sample (see section 2.1.1). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 1 status (variables `kUserGroup` and `kQuitGroup`), and that age at wave 1 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Since longitudinal weights are computed based on the respondents' wave 1 status and data, it is possible for some respondents to be eligible at wave 1, but no longer eligible at wave 2. In such a situation, longitudinal weights are computed for such respondents; whereas no cross-sectional weights are computed, since they are no longer eligible. This typically applies to individuals who are long-term quitters and who do not use NVP (at least monthly). Since respondents who do not use NVP have to had quit smoking within the last 5 years to be eligible, individuals who exceeded the 5 year limit between waves 1 and 2 will have longitudinal weights but no cross-sectional weight.

Though computation followed the same steps as described in section 1.1.2, only 3 Australian exclusive NVP users were successfully recontacted at wave 2 and only 8 in England. Since these are very small samples (and thus their representativeness is likely questionnable), it was decided not to compute waves 1–2 longitudinal weights for those 11 respondents.

1- Variable `lWTS221v` contains the rescaled waves 1–2 longitudinal weights for the 4827 respondents (954 from the US, 1600 from Canada, 1525 from England and 748 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 1 data collection and successfully recontacted at wave 2. These are simply the waves 1–2 longitudinal inflation weights (computed above) of those 4827 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 1 data collection.

2- Variable `lWTS321v` contains the rescaled waves 1–2 longitudinal weights for the 1514 respondents (337 from the US, 572 from Canada, 481 from England and 124 from Australia) who were at least monthly NVP users at the time of wave 1 data collection and successfully recontacted at wave 2. These are simply the waves 1–2 longitudinal inflation weights (computed above) of those 1514 respondents rescaled to sum to sample size in each country. These weights are designed to make these NVP users representative of the population of at least monthly NVP users of their respective country at the time of wave 1 data collection.

   As mentioned above, no waves 1–2 longitudinal weights were computed for British and Australian exclusive NVP users. Hence, for those countries, the `lWTS321v` weights are designed to make these NVP users representative of the population of at least monthly NVP users who are either dual users or quitters using NVP.

3- Variable `lWTS421v` contains the rescaled waves 1–2 longitudinal weights for the 1224 respondents (246 from the US, 487 from Canada, 396 from England and 95 from Australia) who were dual users at the time of wave 1 data collection and successfully recontacted at wave 2. These are simply the waves 1–2 longitudinal inflation weights (computed above) of those 1224 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 1 data collection.

4- Variable `lWTS521v` contains the rescaled waves 1–2 longitudinal weights for the 758 respondents (227 from the US, 235 from Canada, 194 from England and 102 from Australia) who had quit smoking cigarettes at the time of wave 1 data collection and successfully recontacted at wave 2. These are simply the waves 1–2 longitudinal inflation weights (computed above) of those 758 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country at the time of wave 1 data collection. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

5- Variable `lWTS621v` contains the rescaled waves 1–2 longitudinal weights for the 5117 respondents (1045 from the US, 1685 from Canada, 1610 from England and 777 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 1 data collection and successfully recontacted at wave 2. These are simply the waves 1–2 longitudinal inflation weights (computed above) of those 5117 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 1 data collection.

   The note about exclusive NVP users listed for variable `lWTS321v` also applies here.

6- Variable `lWTS121v` contains the rescaled waves 1–2 longitudinal weights for all 5632 respondents (1198 from the US, 1865 from Canada, 1719 from England and 850 from Australia) from the wave 1 main sample that were successfully recontacted at wave 2. These are simply the waves 1–2 longitudinal inflation weights (computed above) of those 5632 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1865 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 1 data collection; likewise for the US, England and Australia (within the last 2 years).

   As mentioned in the description of variable `kWTS101v`, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to

analyse them together using the lWTS121v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be on interest in a longitudinal analysis) from that same joint population is probably much more questionable.

2.2.3 Waves 1–2 longitudinal weights for the US reduced sample

Of the 2239 respondents from the wave 1 US reduced sample, 1129 (or 50.4%) were successfully recontacted and interviewed at wave 2. However, only 12 of those 1129 respondents are exclusive (at least monthly) NVP users. Since such a small sample would have resulted in weights that would have been somewhat unstable, it was decided to exclude those exclusive NVP users, and only compute weights for the remaining 1117 respondents. The following 6 sets of waves 1–2 longitudinal were computed for those respondents. As in section 2.2.2, the following weights are designed to make those 1117 respondents representative at the time of wave 1. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

1- Variable lWTS261v is the analogue of variable lWTS221v, but for the reduced US sample of 906 respondents who were cigarette smokers at the time of wave 1 data collection.

2- Variable lWTS361v is the analogue of variable lWTS321v, but for the reduced US sample of 256 respondents who were at least monthly NVP users at the time of wave 1 data collection.

3- Variable lWTS461v is the analogue of variable lWTS421v, but for the reduced US sample of 198 respondents who were dual users at the time of wave 1 data collection.

4- Variable lWTS561v is the analogue of variable lWTS521v, but for the reduced US sample of 211 respondents who were quitters at the time of wave 1 data collection.

5- Variable lWTS661v is the analogue of variable lWTS621v, but for the reduced US sample of 964 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 1 data collection.

6- Variable lWTS161v is the analogue of variable lWTS121v, but for the reduced US sample of 1117 respondents.

2.2.4 Waves 1–2 longitudinal weights for Australian NVP users

Of the 816 respondents from the wave 1 Australian NVP users sample, 455 (or 55.8%) were successfully recontacted and interviewed at wave 2. The following 3 sets of waves 1–2 longitudinal were computed for those respondents. As in section 2.2.2, the following weights are designed to make those 455 respondents representative at the time of wave 1 data collection. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

1- Variable lWTS381v is the analogue of variable lWTS321v, but for all 455 Australian respondents (124 from the main sample and 331 from the CCV sample; see table 7) who were at least monthly NVP users at the time of wave 1 data collection (and successfully recontacted at wave 2).

2- Variable `lWTS481v` is the analogue of variable `lWTS421v`, but for all 135 Australian respondents (95 from the main sample and 40 from the CCV sample; see table 7) who were dual users at the time of wave 1 data collection (and successfully recontacted at wave 2).

3- Variable `lWTS581v` contains the rescaled waves 1–2 longitudinal weights for all 304 Australian respondents (29 from the main sample and 275 from the CCV sample; see table 7) who had quit cigarette smoking but were using NVP (at least monthly) at the time of wave 1 data collection (and successfully recontacted at wave 2).

3 Wave 3 weights

Eighteen sets of cross-sectional weights and 30 sets of longitudinal weights were computed at wave 3 of the 4CV Survey. These cross-sectional weights are listed in table 8, and their computation and how/when they should be used are detailed in section 3.1. Likewise, the longitudinal weights are listed in table 9, and their computation and use are detailed in sections 3.2 (waves 1–3 longitudinal weights) and 3.3 (waves 2–3 longitudinal weights).

Longitudinal weights for 4C respondents were computed at waves 1 (see section 1.2) and 2 (see section 2.2.1). However, too few 4C respondents were retained at wave 3 for them to be representative on their own, and it was thus decided to no longer compute longitudinal weights for 4C respondents

All sampling weights for the 4CV Survey were computed using the statistical software R (r-project.org). As mentioned at the beginning of this document, these weights adjust for: 1) oversampling of 18–24 years old tobacco users, 2) the oversampling of nicotine vaping product (NVP) users and e-cigarette users, 3) sample mis-representation, 4) non-response and 5) other biases. **It is thus essential to use weighted data, when performing any analyses using 4CV data**.

Weight	Variable Names		
	Main sample⋆	reduced US sample†	AU NVP users‡
Wave 3 cross-sectional inflation weights	mWTS100v	mWTS102v	n/a
Wave 3 cross-sectional inflation weights for NVP users	n/a	n/a	mWTS304v
Rescaled wave 3 cross-sectional weights for cigarette smokers	mWTS201v	mWTS203v	n/a
Rescaled wave 3 cross-sectional weights for NVP users	mWTS301v	mWTS303v	mWTS305v
Rescaled wave 3 cross-sectional weights for dual users	mWTS401v	mWTS403v	mWTS405v
Rescaled wave 3 cross-sectional weights for quitters	mWTS501v	mWTS503v	n/a
Rescaled wave 3 cross-sectional weights for quitters who use NVP	n/a	n/a	mWTS505v
Rescaled wave 3 cross-sectional weights for all tobacco users	mWTS601v	mWTS603v	n/a
Rescaled wave 3 cross-sectional weights for all respondents	mWTS101v	mWTS103v	n/a

⋆ The main sample consists of 11522 respondents (2507 from the US, 3650 from Canada, 3882 from England and 1483 from Australia). It includes all respondents except the 297 Australian respondents from the dedicated CCV sample, the 77 long-term quitters (quit more than 5 years ago: 21 from the US, 38 from Canada and 18 from England) who are not NVP/HTP users, the 64 Australian user who quit more than 2 years ago who are not NVP/HTP users and the 90 respondents (26 from the US, 33 from Canada, 25 from England, and 6 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).
† The reduced sample is the main US sample of 2507 respondents minus the 62 NVP users recruited by Ipsos at wave 1 and successfully recontacted at wave 3; hence, the reduced US sample consists of 2445 respondents.
‡ This sample contains all 586 Australian respondents that were vaping at the time of wave 3. As detailed in section 3.1.4, this was done to allow for weighted analyses including the 297 respondents recruited via the dedicated CCV sample.

Table 8: Cross-sectional sampling weights for wave 3 of the 4CV Survey

Weight	Variable Names		
	Main sample	reduced US sample	AU NVP users
Rescaled waves 1–3 longitudinal weights for cigarette smokers	mWTS221v	mWTS261v	n/a
Rescaled waves 1–3 longitudinal weights for NVP users	mWTS321v	mWTS361v	mWTS381v
Rescaled waves 1–3 longitudinal weights for dual users	mWTS421v	mWTS461v	mWTS481v
Rescaled waves 1–3 longitudinal weights for quitters	mWTS521v	mWTS561v	n/a
Rescaled waves 1–3 longitudinal weights for quitters who use NVP	n/a	n/a	mWTS581v
Rescaled waves 1–3 longitudinal weights for all tobacco users	mWTS621v	mWTS661v	n/a
Rescaled waves 1–3 longitudinal weights for all respondents	mWTS121v	mWTS161v	n/a
Rescaled waves 2–3 longitudinal weights for cigarette smokers	mWTS223v	mWTS263v	n/a
Rescaled waves 2–3 longitudinal weights for NVP users	mWTS323v	mWTS363v	mWTS383v
Rescaled waves 2–3 longitudinal weights for dual users	mWTS423v	mWTS463v	mWTS483v
Rescaled waves 2–3 longitudinal weights for quitters	mWTS523v	mWTS563v	n/a
Rescaled waves 2–3 longitudinal weights for quitters who use NVP	n/a	n/a	mWTS583v
Rescaled waves 2–3 longitudinal weights for all tobacco users	mWTS623v	mWTS663v	n/a
Rescaled waves 2–3 longitudinal weights for all respondents	mWTS123v	mWTS163v	n/a

Table 9: Longitudinal sampling weights for wave 3 of the 4CV Survey

3.1 Cross-sectional sampling weights

As with previous waves, the 18 sets of cross-sectional weights can be divided into 3 categories:

1. The first 7 sets of cross-sectional weights (see section 3.1.2 and column 2 of table 8) were computed for the main sample. That sample consists of 11522 respondents (2507 from the US, 3650 from Canada, 3882 from England and 1483 from Australia; see table 10). It includes all respondents except the 297 Australian respondents from the dedicated CCV sample, the 141 long-term quitters (21 from the US, 38 from Canada, 18 from England, and 64 from Australia) who are not NVP/HTP users, and the 90 respondents (26 from the US, 33 from Canada, 25 from England, and 6 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).

2. The next 7 sets of cross-sectional weights (see section 3.1.3 and column 3 of table 8) were computed for the reduced US sample. This sample is the main US sample of 2507 respondents (mentioned above) minus the 62 NVP users recruited via Ipsos' non-probability Web panel at wave 1 and successfully recontacted at waves 2 and 3; hence, the reduced US sample consists of 2445 respondents. Those 62 respondents are the only US respondents not recruited using a probability based method, and those sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired. Since this only applies to US respondents, no such weights were computed for respondents from Canada, England or Australia.

3. The last 3 sets of cross-sectional weights (see section 3.1.4 and column 4 of table 8) were computed for Australian respondents using NVP at least monthly. This sample contains all 586 Australian respondents (289 from the main sample and 297 from the CCV sample) that were vaping at the time of wave 3. This was done to allow for weighted analyses including the 297 respondents recruited via the dedicated CCV sample. This concerns Australian respondents only, and no such weights were computed for respondents from the US, Canada or England.

3.1.1 User groups

As in previous waves, the first step in the computation of all sets of wave 3 cross-sectional weights consisted in dividing respondents into 4 user groups (variable `mUserGroup` in the dataset), and then further dividing quitters into 4 sub-groups (variable `mQuitGroup` in the dataset). However, how those groups and sub-groups are defined was slightly updated at wave 3. This was done to include users of heated tobacco products (HTPs), and Australians who have quit smoking cigarettes within the last 2–5 years (as opposed to only within the last 2 years). Hence, starting at wave 3, the user group definitions of section 1.1.1 were updated to the following:

- To be classified as a cigarette only smoker (group *i*), a respondent had to (at the time of data collection) smoke cigarettes (at least occasionally), smoked at least 100 cigarettes in their lifetime, and use NVP/HTP (this includes e-cigarettes) less than monthly or not at all.

  Note that the comment made in section 1.1.1 about including those less than monthly but occasional smokers also applies here, and below.

- To be classified as a dual user (group *ii*), a respondent had to (at the time of data collection) smoke cigarettes (at least occasionally), smoked at least 100 cigarettes in their lifetime, and use NVP/HTP (this includes e-cigarettes) at least monthly.

  Note that the comment made in section 1.1.1 about the use of the at "least monthly" criteria for the use of NVP being a limitation of our study also applies here, and below.

- To be classified as an exclusive NVP/HTP user (group *iii*), a respondent had to (at the time of data collection) use NVP/HTP (this includes e-cigarettes) at least monthly, and satisfy one or more of the following:
  - not currently smoking cigarettes
  - smoked less than 100 cigarettes in their lifetime
  - quit smoking cigarettes more than 5 years ago
- To be classified as a quitter (group *iv*), a respondent had to (at the time of data collection) have quit smoking cigarettes within the last 5 years, and smoked at least 100 cigarettes in their lifetime prior to that. Those who were using NVP/HTP (this includes e-cigarettes) at least monthly were then further sub-classified in either group *iv.a* or *iv.b* depending on how long ago they had quit smoking; whereas those who were using NVP/HTP less than monthly or not at all were then further sub-classified in either group *iv.c* or *iv.d* depending on how long ago they had quit smoking.

The numbers of respondents in each of those user groups are given in table 10. It should be noted that, except for the British 2019 Opinions and Lifestyle Survey (OLS), the national health surveys used to calibrate the wave 3 cross-sectional weights provided little information to properly estimate the number of HTP users. Hence, the benchmark figures for user groups *ii*, *iii*, *iv.a* and *iv.b* are likely to somewhat underestimate the true values in Canada, the US and Australia. Therefore, the weights for those user groups are likely to be slightly too small. It should additionally be noted that, in our wave 3 sample, many HTP users are also NVP users. Consequently, the above grouping where NVP and HTP are "merge" adjusts for the oversampling of those users via the dedicated NVP/HTP/e-cigarettes samples.

User group†	US		Canada	England	Australia	
	Main‡	Reduced‡			Main‡	NVP‡
Cigarette smokers						
Cigarette only	1275	1244	1802	1857	1048	0
Dual users	580	569	1112	1429	183	192
Total	1855	1813	2914	3286	1231	192
Exclusive NVP/HTP users	256	251	171	198	49	163
Quitters						
NVP/HTP users	143	138	143	168	57	231
Non users	253	243	422	230	146	0
Total	396	381	565	398	203	231
Total	2507	2445	3650	3882	1483	586

† Variables `mUserGroup` and `mQuitGroup` in the dataset
‡ See notes below table 8 and beginning of section 3.1

Table 10: Wave 3 respondents by country and user group.

3.1.2 Cross-sectional weights for the main sample

1- Variable `mWTS100v` contains the wave 3 cross-sectional inflation weights for the main sample of 11522 respondents (2507 from the US, 3650 from Canada, 3882 from England and 1483 from Australia; see

table 10). Computation of those weights followed the same steps and used the same raking algorithm (see appendix A.3) as the wave 1 cross-sectional inflation weights for the main sample (variable `kWTS100v` of section 1.1.2). The only differences being:

i) The wave 3 cross-sectional inflation weights are designed to make the sample representative of the target population at the time of wave 3. Consequently, for cohort respondents, age at wave 3 was used instead of their age when they were first recruited. Likewise, variables `mUserGroup` and `mQuitGroup` (as updated in section 3.1.1) were used instead of `kUserGroup` and `kQuitGroup`, and similarly for education and geographic region. It is thus possible for a respondent to transition from being a cigarette only smoker (group *i*) at wave 1 and to being classified as a dual user (group *ii*) at wave 2. Other types of transitions are also possible. Likewise, it is possible for respondents to move between geographic regions or increase his/her education level.

ii) Except for user group × language in Canada, the same cross-tabs were used (i.e., user group × gender, user group × age group, user group × geographic region, user group × ethnicity (US only), and user group × education (except for Canada)). Language was not used to calibrate the Canadian wave 3 cross-sectional weights, as this information was unfortunately not available in the benchmark survey. In addition, the collapsing of cells (in particular exclusive NVP/HTP users), because they contained too few respondents, was slightly different at wave 3 than at prior waves. To be precise, Canadian and American exclusive NVP/HTP users (group *iii*) were only further subdivided based on gender and age (18–24 vs. > 24); whereas, the 198 British and 49 Australian respondents in group *iii* were not further subdivided.

iii) As with the wave 2 cross-sectional weights (e.g., variable `1WTS100v`), updated benchmark/calibration figures were used to calibrate the wave 3 cross-sectional inflation weights. In the US, the 2018 National Health Interview Survey (NHIS) was used; whereas the 2019 National Drug Strategy Household Survey (NDSHS) was used in Australia. In England, the 2019 Opinions and Lifestyle Survey (OLS) was combined with waves 161–164 (Feb–Jun 2020) of the Smoking Toolkit Study (STS). Hence, in those 3 countries, more recent iterations of the same national health surveys (used for waves 1 and 2) were used to calibrate the wave 3 cross-sectional inflation weights. In Canada, the Canadian Tobacco Alcohol and Drugs Survey (CTADS), used to calibrate the weights at waves 1 and 2, was replaced by the Canadian Tobacco and Nicotine Survey (CTNS). Unfortunately, the 2019 CTNS did not provide all the information required to compute the benchmark/calibration figures for individuals aged 18–19 years old. Hence, the 2017 CTADS (used at wave 2) was combined with the 2019 CTNS for ages 18–19; whereas, benchmark/calibration figures for individuals aged 20 & older are solely based on the 2019 CTNS. See the "Wave 3" column of the various tables in appendix A.2.

These weights are designed to make respondents in each of the four groups representative of the corresponding population at the time of wave 3 data collection. For example, the `mWTS100v` weights of the 1112 Canadian dual users are designed to make them representative of the Canadian population of dual users at the time of data collection; likewise for the other countries and the other groups. If interests lie in a target population that consists of two or more of the four user groups, the `mWTS100v` weights are still appropriate. For example, when studying Canadian cigarette smokers, one can simply combine the `mWTS100v` weights of the 1802 cigarette only users with those of the 1112 dual users (for a total of 2914 respondents in the analysis), and assign a weight of 0 to respondents in the other two user groups.

Last but not least, since these are inflation/un-rescaled weights, they should not be used in analyses involving two or more countries. The various rescaled weights (i.e., variables `mWTS101v` to `mWTS601v`)

described below were created especially for such multi-country analyses; see section 6.2 for more information on inflation versus rescaled weights.

2- Variable mWTS201v contains the rescaled wave 3 cross-sectional weights for the 9286 (1855 from the US, 2914 from Canada, 3286 from England and 1231 from Australia; see table 10) respondents who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS100v) of those 9286 respondents rescaled to sum to sample size in each country (i.e., 1855 in the US, 2914 in Canada, 3286 in England and 1231 in Australia). These weights are designed to make these 2914 Canadian cigarette smokers representative of the Canadian population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 3 data collection; likewise for the US, England and Australia.

3- Variable mWTS301v contains the rescaled wave 3 cross-sectional weights for the 4489 (979 from the US, 1426 from Canada, 1795 from England and 289 from Australia; see table 10) respondents who were at least monthly NVP/HTP users at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS100v) of those 4489 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1426 Canadian NVP/HTP users representative of the Canadian population of at least monthly NVP/HTP users at the time of wave 3 data collection; likewise for the US, England and Australia.

4- Variable mWTS401v contains the rescaled wave 3 cross-sectional weights for the 3304 (580 from the US, 1112 from Canada, 1429 from England and 183 from Australia; see table 10) respondents who were dual users at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS100v) of those 3304 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1112 Canadian dual users representative of the Canadian population of dual users at the time of wave 3 data collection; likewise for the US, England and Australia.

5- Variable mWTS501v contains the rescaled wave 3 cross-sectional weights for the 1562 (396 from the US, 565 from Canada, 398 from England and 203 from Australia; see table 10) quitters (within the last 5 years) at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS100v) of those 1562 respondents rescaled to sum to sample size in each country. The cautionary notes of variable kWTS501v and of section 6.3 still apply here.

6- Variable mWTS601v contains the rescaled wave 3 cross-sectional weights for the 10471 (2254 from the US, 3228 from Canada, 3652 from England and 1337 from Australia; see table 10) respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users[11]) at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS100v) of those 10471 respondents rescaled to sum to sample size in each country. These weights are designed to make these 3228 Canadian tobacco users representative of the Canadian population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users) at the time of wave 3 data collection; likewise for the US, England and Australia.

7- Variable mWTS101v contains the rescaled wave 3 cross-sectional weights for the main sample of 11522 respondents (2507 from the US, 3650 from Canada, 3882 from England and 1483 from Australia; see table 10). These are simply the wave 3 cross-sectional inflation weights (variable mWTS100v) of

[11] Note that some of those NVP/HTP users are individuals who have quit smoking traditional cigarette.

those 11522 respondents rescaled to sum to sample size in each country (i.e., 2507 in the US, 3650 in Canada, 3882 in England and 1483 in Australia). These weights are designed to make these 3650 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters at the time of wave 3 data collection; likewise for the US, England and Australia.

As mentioned in the description of variable kWTS101v, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the mWTS101v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (e.g., linear regression and logistic regression) from that same joint population is probably much more questionable.

3.1.3 Cross-sectional weights for the reduced US sample

As mentioned at the beginning of section 3.1, the following 7 sets of weights only apply to the US sample.

1- Variable mWTS102v contains the wave 3 cross-sectional inflation weights for the reduced US sample of 2445 respondents. As noted below table 8, the reduced sample is the main sample of 2507 respondents minus the 62 NVP users recruited by Ipsos at wave 1 (see section 1.1.3), and retained at waves 2 and 3. Those NVP users are identifiable via the variable mOWNERID[12] in the dataset. Those at least monthly NVP users were recruited via a non-probability based panel, and the 7 sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired for their specific analyses.

   These weights were computed the same way as the wave 3 cross-sectional inflation weights for the main sample (see variable mWTS100v in section 3.1.2).

2- Variable mWTS203v contains the rescaled wave 3 cross-sectional weights for the 1813 respondents (see table 10) from the reduced US sample who were cigarette smokers at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS102v) of those respondents rescaled to sum to sample size (i.e, 1813). These weights are designed to make these 1813 US cigarette smokers representative of the US population of cigarette smokers at the time of wave 3 data collection. In other words, variable mWTS203v is the analogue of variable mWTS201v, but for the reduced sample.

3- Variable mWTS303v is the analogue of variable mWTS301v, but for the reduced US sample of 958 respondents who were at least monthly NVP/HTP users at the time of wave 3 data collection.

4- Variable mWTS403v is the analogue of variable mWTS401v, but for the reduced US sample of 569 respondents who were dual users at the time of wave 3 data collection.

5- Variable mWTS503v is the analogue of variable mWTS501v, but for the reduced US sample of 381 respondents who were quitters at the time of wave 3 data collection.

6- Variable mWTS603v is the analogue of variable mWTS601v, but for the reduced US sample of 2202 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users) at the time of wave 3 data collection.

[12] Use mOWNERID!=05 to exclude the 62 US NVP users recruited by Ipsos, where != means not equal to.

7- Variable mWTS103v is the analogue of variable mWTS101v, but for the reduced US sample of 2445 respondents. Hence, the cautionary notes listed for variable mWTS101v also apply here.

3.1.4 Cross-sectional weights for Australian NVP users

As mentioned at the beginning of section 3.1, the following 3 sets of weights only apply to the Australian sample of 586 respondents (289 from the main sample and 297 from the CCV sample; see table 10) that were using NVP/HTP (at least monthly) at the time of wave 3. No such weights were computed for respondents from the US, Canada, England or for Australian respondents that are not at least monthly NVP/HTP users.

Recall from section 1.1.4 that the CCV vaper sample (identifiable via the variable mOWNERID[13] in the dataset) consists of self-selected individuals recruited mainly through online sites. They are a group of vaper activists and early adopters. They are thus not a representative sample of vaper users in Australia. This is why those respondents are excluded from the 7 sets of weights computed in section 3.1.2. Nevertheless, respondents from the CCV sample can be of scientific interest, and some data users might want to include them in their analyses. This is why the following 4 sets of weights were computed.

1- Variable mWTS304v contains the wave 3 cross-sectional inflation weights for the Australian sample of 586 respondents (289 from the main sample and 297 from the CCV sample; see table 10) that were using NVP/HTP at least monthly at the time of wave 3.

   These weights were computed in a very similar way as the wave 3 cross-sectional inflation weights for the main sample (see variable mWTS100v in section 3.1.2). However, respondents were first divided into 3 (dual users, exclusive NVP/HTP users and quitters using NVP/HTP) instead of 4 user groups. The raking algorithm was then used to calibrate the weights on user group × gender, user group × age group, and the 2 quitter sub-groups (i.e., quit within the last year and use NVP/HTP, and quit more than 1 year ago and use NVP/HTP).

2- Variable mWTS305v contains the rescaled wave 3 cross-sectional weights for the 586 Australian respondents who were at least monthly NVP/HTP users at the time of wave 3 data collection. These are simply the wave 3 cross-sectional inflation weights (variable mWTS304v) of those respondents rescaled to sum to sample size (i.e, 586). These weights are designed to make these 586 NVP/HTP users representative of the Australian population of at least monthly NVP/HTP users at the time of wave 3 data collection. In other words, variable mWTS305v is the analogue of variable mWTS301v, but for all 586 Australian at least monthly NVP/HTP users.

3- Variable mWTS405v is the analogue of variable mWTS401v, but for the 192 (183 from the main sample and 9 from the CCV sample; see table 10) Australian respondents who were dual users at the time of wave 3 data collection.

4- Variable mWTS505v contains the rescaled wave 3 cross-sectional weights for the 231 (57 from the main sample and 174 from the CCV sample; see table 10) Australian respondents who had quit cigarette smoking but were using NVP/HTP at least monthly at the time of wave 3 data collection.

[13] Use mOWNERID $\in \{07, 70\}$ to select the 297 respondents from the dedicated CCV sample.

3.2 Waves 1–3 longitudinal sampling weights

As mentioned at the beginning of section 3 (table 9), 30 sets of longitudinal weights were computed at wave 3 of the 4CV Survey. Half of those (see top half of table 9) were computed for respondents recruited at wave 1, and then successfully retained and interviewed at waves 2 and 3. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 1—3 longitudinal weights are designed to make the sample representative at the time of wave 1. If the aim is to ensure that the sample is representative at the time of wave 3 data collection, then the wave 3 cross-sectional sampling weights computed in section 3.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

User group†	US		Canada	England	Australia	
	Main	Reduced			Main	NVP
Cigarette only	463	463	636	669	438	0
Dual users	158	140	251	212	64	93
Exclusive NVP users	0	0	15	0	0	9
Quitters						
NVP users	51	41	36	60	21	179
Non users	27	27	74	46	34	0
Total	78	68	110	106	55	179
Total	699	671	1012	987	557	281

† Variables `kUserGroup` and `kQuitGroup` in the dataset

Table 11: Wave 1 respondents successfully recontacted at wave 3 by user group.

3.2.1 Waves 1–3 longitudinal weights for the main sample

Waves 1–3 longitudinal inflation weights were computed for the 3255 respondents (699 from the US, 1012 from Canada, 987 from England and 557 from Australia) from the wave 1 main sample that were successfully recontacted at both waves 2 and 3. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 1–3 longitudinal inflation weights are the waves 1–2 longitudinal inflation weights (see section 2.2.2) adjusted for attrition between waves 2 and 3. As with the waves 1–2 longitudinal inflation weights, computation of the waves 1–3 longitudinal inflation weights followed the same steps as described in section 1.1.2. Since these weights are meant to make these 3255 respondents representative of the population at the time of wave 1 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 1 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 3 cross-sectional weights for the main sample (see section 3.1.2). Note that respondents were divided into the 4 user groups and 4 quitter subgroups

based on their wave 1 status (variables `kUserGroup` and `kQuitGroup`), and that age at wave 1 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Though computation followed the same steps as described in section 1.1.2, only 9 American and 6 British exclusive NVP users were successfully recontacted at wave 3. Since these are very small samples (and thus their representativeness is likely questionnable), it was decided not to compute waves 1–3 longitudinal weights for those respondents. Furthermore, no Australian exclusive NVP users were successfully recontacted at wave 3, and thus the only exclusive NVP users for which waves 1–3 longitudinal weights were computed are Canadians.

1- Variable `mWTS221v` contains the rescaled waves 1–3 longitudinal weights for the 2891 respondents (621 from the US, 887 from Canada, 881 from England and 502 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 1 data collection and successfully recontacted at wave 3. These are simply the waves 1–3 longitudinal inflation weights (computed above) of those 2891 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 1 data collection.

2- Variable `mWTS321v` contains the rescaled waves 1–3 longitudinal weights for the 868 respondents (209 from the US, 302 from Canada, 272 from England and 85 from Australia) who were at least monthly NVP users at the time of wave 1 data collection and successfully recontacted at wave 3. These are simply the waves 1–3 longitudinal inflation weights (computed above) of those 868 respondents rescaled to sum to sample size in each country. These weights are designed to make these NVP users representative of the population of at least monthly NVP users of their respective country at the time of wave 1 data collection.

   As mentioned above, no waves 1–3 longitudinal weights were computed for American, British and Australian exclusive NVP users. Hence, for those countries, the `mWTS321v` weights are designed to make these NVP users representative of the population of at least monthly NVP users who are either dual users or quitters using NVP.

3- Variable `mWTS421v` contains the rescaled waves 1–3 longitudinal weights for the 685 respondents (158 from the US, 251 from Canada, 212 from England and 64 from Australia) who were dual users at the time of wave 1 data collection and successfully recontacted at wave 3. These are simply the waves 1–3 longitudinal inflation weights (computed above) of those 685 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 1 data collection.

4- Variable `mWTS521v` contains the rescaled waves 1–3 longitudinal weights for the 349 respondents (78 from the US, 110 from Canada, 106 from England and 55 from Australia) who had quit smoking cigarettes at the time of wave 1 data collection and successfully recontacted at wave 3. These are simply the waves 1–3 longitudinal inflation weights (computed above) of those 349 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country at the time of wave 1 data collection. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

5- Variable `mWTS621v` contains the rescaled waves 1–3 longitudinal weights for the 3074 respondents (672 from the US, 938 from Canada, 941 from England and 523 from Australia) who were tobacco users

(i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 1 data collection and successfully recontacted at wave 3. These are simply the waves 1–3 longitudinal inflation weights (computed above) of those 3074 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 1 data collection.

The note about exclusive NVP users listed for variable mWTS321v also applies here.

6- Variable mWTS121v contains the rescaled waves 1–3 longitudinal weights for all 3255 respondents (699 from the US, 1012 from Canada, 987 from England and 557 from Australia) from the wave 1 main sample that were successfully recontacted at wave 3. These are simply the waves 1–3 longitudinal inflation weights (computed above) of those 3255 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1012 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 1 data collection; likewise for the US, England and Australia (within the last 2 years).

As mentioned in the description of variable kWTS101v, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the mWTS121v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

3.2.2 Waves 1–3 longitudinal weights for the US reduced sample

The following 6 sets of waves 1–3 longitudinal weights were computed for the 671 respondents from the wave 1 US reduced sample (see third column of table 11) that were successfully recontacted at both waves 2 and 3. Computation followed the same steps as described in section 3.2.1, and the following weights are also designed to make those 671 respondents representative at the time of wave 1. Note that none of the exclusive NVP users from the wave 1 reduced sample were successfully recontacted at wave 3. Hence, the following weights are designed to make these NVP users representative of the population of at least monthly NVP users who are either dual users or quitters using NVP.

1- Variable mWTS261v is the analogue of variable mWTS221v, but for the reduced US sample of 603 respondents who were cigarette smokers at the time of wave 1 data collection.

2- Variable mWTS361v is the analogue of variable mWTS321v, but for the reduced US sample of 181 respondents who were at least monthly NVP users at the time of wave 1 data collection.

The note about exclusive NVP users listed for variable mWTS321v also applies here.

3- Variable mWTS461v is the analogue of variable mWTS421v, but for the reduced US sample of 140 respondents who were dual users at the time of wave 1 data collection.

4- Variable mWTS561v is the analogue of variable mWTS521v, but for the reduced US sample of 68 respondents who were quitters at the time of wave 1 data collection.

5- Variable mWTS661v is the analogue of variable mWTS621v, but for the reduced US sample of 644 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 1 data collection.

   The note about exclusive NVP users listed for variable mWTS321v also applies here.

6- Variable mWTS161v is the analogue of variable mWTS121v, but for the reduced US sample of 671 respondents.

3.2.3 Waves 1–3 longitudinal weights for Australian NVP users

The following 3 sets of waves 1–3 longitudinal weights were computed for the 281 respondents from the wave 1 Australian NVP users sample (88 from the main sample and 193 from the CCV sample; see last column of table 11) that were successfully recontacted at both waves 2 and 3. Computation followed steps similar to those detailed in section 3.2.1, and the following weights are also designed to make those 281 respondents representative at the time of wave 1.

1- Variable mWTS381v is the analogue of variable mWTS321v, but for all 281 Australian respondents (88 from the main sample and 193 from the CCV sample; see table 11) who were at least monthly NVP users at the time of wave 1 data collection (and successfully recontacted at wave 3).

2- Variable mWTS481v is the analogue of variable mWTS421v, but for all 93 Australian respondents (64 from the main sample and 29 from the CCV sample; see table 11) who were dual users at the time of wave 1 data collection (and successfully recontacted at wave 3).

3- Variable mWTS581v contains the rescaled waves 1–3 longitudinal weights for all 179 Australian respondents (21 from the main sample and 158 from the CCV sample; see table 11) who had quit cigarette smoking but were using NVP (at least monthly) at the time of wave 1 data collection (and successfully recontacted at wave 3).

3.3 Waves 2–3 longitudinal sampling weights

As mentioned at the beginning of section 3, 30 sets of longitudinal weights were computed at wave 3 of the 4CV Survey. Half of those (see bottom half of table 9) were computed for respondents recruited at waves 1 and 2, and then successfully retained and interviewed at wave 3. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 2—3 longitudinal weights are designed to make the sample representative at the time of wave 2. If the aim is to ensure that the sample is representative at the time of wave 3 data collection, then the wave 3 cross-sectional sampling weights computed in section 3.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

3.3.1 Waves 2–3 longitudinal weights for the main sample

Waves 2–3 longitudinal inflation weights were computed for the 5475 respondents (1363 from the US, 1658 from Canada, 1586 from England and 868 from Australia) from the wave 2 main sample that were successfully recontacted at wave 3. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled

User group†	US		Canada	England	Australia	
	Main	Reduced			Main	NVP
Cigarette only	777	746	932	881	626	0
Dual users	264	249	378	443	100	119
Exclusive NVP users	45	41	12	18	0	41
Quitters						
NVP users	125	119	81	124	45	316
Non users	152	144	255	120	97	0
Total	277	263	336	244	142	316
Total	1363	1299	1658	1586	868	476

† Variables `lUserGroup` and `lQuitGroup` in the dataset

Table 12: Wave 2 respondents successfully recontacted at wave 3 by user group.

weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 2–3 longitudinal inflation weights are the wave 2 cross-sectional inflation weights (i.e., variable `lWTS100v`, computed in section 2.1.1) adjusted for attrition between waves 2 and 3. Consequently, computation of those longitudinal inflation weights followed the same steps as described in section 2.1.1. Since these weights are meant to make these 5475 respondents representative of the population at the time of wave 2 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 2 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 3 cross-sectional weights for the main sample (see section 3.1.2). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 2 status (variables `lUserGroup` and `lQuitGroup`), and that age at wave 2 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Though computation followed the same steps as described in section 2.1.1, only 7 Australian exclusive NVP users were successfully recontacted at wave 3. Since this is a very small sample (and thus its representativeness is likely questionnable), it was decided not to compute waves 2–3 longitudinal weights for those respondents.

1- Variable `mWTS223v` contains the rescaled waves 2–3 longitudinal weights for the 4401 respondents (1041 from the US, 1310 from Canada, 1324 from England and 726 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 2 data collection and successfully recontacted at wave 3. These are simply the waves 2–3 longitudinal inflation weights (computed above) of those 4401 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 2 data collection.

2- Variable `mWTS323v` contains the rescaled waves 2–3 longitudinal weights for the 1635 respondents (434 from the US, 471 from Canada, 585 from England and 145 from Australia) who were at least monthly NVP users at the time of wave 2 data collection and successfully recontacted at wave 3. These are

simply the waves 2–3 longitudinal inflation weights (computed above) of those 1635 respondents rescaled to sum to sample size in each country. These weights are designed to make these NVP users representative of the population of at least monthly NVP users of their respective country at the time of wave 2 data collection.

As mentioned above, no waves 2–3 longitudinal weights were computed for Australian exclusive NVP users. Hence, for that country, the `mWTS323v` weights are designed to make these NVP users representative of the population of at least monthly NVP users who are either dual users or quitters using NVP.

3- Variable `mWTS423v` contains the rescaled waves 2–3 longitudinal weights for the 1185 respondents (264 from the US, 378 from Canada, 443 from England and 100 from Australia) who were dual users at the time of wave 2 data collection and successfully recontacted at wave 3. These are simply the waves 2–3 longitudinal inflation weights (computed above) of those 1185 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 2 data collection.

4- Variable `mWTS523v` contains the rescaled waves 2–3 longitudinal weights for the 999 respondents (277 from the US, 336 from Canada, 244 from England and 142 from Australia) who had quit smoking cigarettes at the time of wave 2 data collection and successfully recontacted at wave 3. These are simply the waves 2–3 longitudinal inflation weights (computed above) of those 999 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country at the time of wave 2 data collection. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

5- Variable `mWTS623v` contains the rescaled waves 2–3 longitudinal weights for the 4851 respondents (1211 from the US, 1403 from Canada, 1466 from England and 771 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection and successfully recontacted at wave 3. These are simply the waves 2–3 longitudinal inflation weights (computed above) of those 4851 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 2 data collection.

The note about exclusive NVP users listed for variable `mWTS323v` also applies here.

6- Variable `mWTS123v` contains the rescaled waves 2–3 longitudinal weights for all 5475 respondents (1363 from the US, 1658 from Canada, 1586 from England and 868 from Australia) from the wave 2 main sample that were successfully recontacted at wave 3. These are simply the waves 2–3 longitudinal inflation weights (computed above) of those 5475 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1658 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 2 data collection; likewise for the US, England and Australia (within the last 2 years).

As mentioned in the description of variable `mWTS101v`, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the `mWTS123v` weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

3.3.2 Waves 2–3 longitudinal weights for the US reduced sample

The following 6 sets of waves 2–3 longitudinal weights were computed for the 1299 respondents from the wave 2 US reduced sample (see third column of table 12) that were successfully recontacted at waves 3. Computation followed the same steps as described in section 3.3.1, and the following weights are also designed to make those 1299 respondents representative at the time of wave 2.

1- Variable `mWTS263v` is the analogue of variable `mWTS223v`, but for the reduced US sample of 995 respondents who were cigarette smokers at the time of wave 2 data collection.

2- Variable `mWTS363v` is the analogue of variable `mWTS323v`, but for the reduced US sample of 409 respondents who were at least monthly NVP users at the time of wave 2 data collection.

3- Variable `mWTS463v` is the analogue of variable `mWTS423v`, but for the reduced US sample of 249 respondents who were dual users at the time of wave 2 data collection.

4- Variable `mWTS563v` is the analogue of variable `mWTS523v`, but for the reduced US sample of 263 respondents who were quitters at the time of wave 1 data collection.

5- Variable `mWTS663v` is the analogue of variable `mWTS623v`, but for the reduced US sample of 1155 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection.

6- Variable `mWTS163v` is the analogue of variable `mWTS123v`, but for the reduced US sample of 1299 respondents.

3.3.3 Waves 2–3 longitudinal weights for Australian NVP users

The following 3 sets of waves 2–3 longitudinal weights were computed for the 476 respondents from the wave 2 Australian NVP users sample (152 from the main sample and 324 from the CCV sample; see last column of table 12) that were successfully recontacted at waves 3. Computation followed steps similar to those detailed in section 3.3.1, and the following weights are also designed to make those 152 respondents representative at the time of wave 2.

1- Variable `mWTS383v` is the analogue of variable `mWTS323v`, but for all 476 Australian respondents (152 from the main sample and 324 from the CCV sample; see table 12) who were at least monthly NVP users at the time of wave 2 data collection (and successfully recontacted at wave 3).

2- Variable `mWTS483v` is the analogue of variable `mWTS423v`, but for all 119 Australian respondents (100 from the main sample and 19 from the CCV sample; see table 12) who were dual users at the time of wave 2 data collection (and successfully recontacted at wave 3).

3- Variable `mWTS583v` contains the rescaled waves 2–3 longitudinal weights for all 316 Australian respondents (45 from the main sample and 271 from the CCV sample; see table 12) who had quit cigarette smoking but were using NVP (at least monthly) at the time of wave 2 data collection (and successfully recontacted at wave 3).

4 Wave 4 weights

Fourteen sets of cross-sectional weights and 30 sets of longitudinal weights were computed at wave 4 of the 4CV Survey. These cross-sectional weights are listed in table 13, and their computation and how/when they should be used are detailed in section 4.1. Likewise, the longitudinal weights are listed in table 14, and their computation and use are detailed in sections 4.2 (waves 1–4 longitudinal weights), 4.3 (waves 2–4 longitudinal weights) and 4.4 (waves 3–4 longitudinal weights).

As with wave 3, too few 4C respondents were retained at wave 4 for them to be representative on their own. It was thus decided to no longer compute longitudinal weights going back to waves 8, 9, and 10 of the 4C Survey (see section 1.2 and 2.2.1 for which weights are computed at waves 1 and 2 of the 4CV Survey). However, 4C respondents are included in the various cross-sectional and longitudinal weights computed in section 4.1–4.4.

All sampling weights for the 4CV Survey were computed using the statistical software R (r-project.org). As mentioned at the beginning of this document, these weights adjust for: 1) oversampling of 18–24 years old tobacco users (at waves 1–3 in US, Canada & England), 2) the oversampling of nicotine vaping product (NVP) users and e-cigarette users[14], 3) sample mis-representation, 4) non-response and 5) other biases. **It is thus essential to use weighted data, when performing any analyses using 4CV data**.

	Variable Names	
Weight	Main sample⋆	reduced US sample†
Wave 4 cross-sectional inflation weights	`nWTS100v`	`nWTS102v`
Rescaled wave 4 cross-sectional weights for cigarette smokers	`nWTS201v`	`nWTS203v`
Rescaled wave 4 cross-sectional weights for NVP users	`nWTS301v`	`nWTS303v`
Rescaled wave 4 cross-sectional weights for dual users	`nWTS401v`	`nWTS403v`
Rescaled wave 4 cross-sectional weights for quitters	`nWTS501v`	`nWTS503v`
Rescaled wave 4 cross-sectional weights for all tobacco users	`nWTS601v`	`nWTS603v`
Rescaled wave 4 cross-sectional weights for all respondents	`nWTS101v`	`nWTS103v`

⋆ The main sample consists of 7991 respondents (2107 from the US, 2120 from Canada, 2267 from England and 1497 from Australia). It includes all respondents except 210 long-term quitters (quit more than 5 years ago: 77 from the US, 64 from Canada, 45 from England and 24 from Australia) who are not NVP/HTP users, and the 71 respondents (21 from the US, 24 from Canada, 21 from England, and 5 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).
† The reduced sample is the main US sample of 2107 respondents minus the 30 NVP users recruited by Ipsos at wave 1 and successfully recontacted at wave 4; hence, the reduced US sample consists of 2077 respondents.

Table 13: Cross-sectional sampling weights for wave 4 of the 4CV Survey

4.1 Cross-sectional sampling weights

Similarly to previous waves, the 14 sets of cross-sectional weights can be divided into 2 categories:

[14] At all waves of US, Canada & England, and waves 1–3 of Australia

Weight	Variable Names	
	Main sample	reduced US sample
Rescaled waves 1–4 longitudinal weights for cigarette smokers	nWTS221v	nWTS261v
Rescaled waves 1–4 longitudinal weights for NVP users	n/a†	n/a†
Rescaled waves 1–4 longitudinal weights for dual users	nWTS421v	nWTS461v
Rescaled waves 1–4 longitudinal weights for quitters	n/a†	n/a†
Rescaled waves 1–4 longitudinal weights for all tobacco users	n/a†	n/a†
Rescaled waves 1–4 longitudinal weights for all respondents	nWTS121v	nWTS161v
Rescaled waves 2–4 longitudinal weights for cigarette smokers	nWTS223v	nWTS263v
Rescaled waves 2–4 longitudinal weights for NVP users	nWTS323v‡	nWTS363v
Rescaled waves 2–4 longitudinal weights for dual users	nWTS423v	nWTS463v
Rescaled waves 2–4 longitudinal weights for quitters	nWTS523v‡	nWTS563v
Rescaled waves 2–4 longitudinal weights for all tobacco users	nWTS623v	nWTS663v
Rescaled waves 2–4 longitudinal weights for all respondents	nWTS123v	nWTS163v
Rescaled waves 3–4 longitudinal weights for cigarette smokers	nWTS225v	nWTS265v
Rescaled waves 3–4 longitudinal weights for NVP users	nWTS325v	nWTS365v
Rescaled waves 3–4 longitudinal weights for dual users	nWTS425v	nWTS465v
Rescaled waves 3–4 longitudinal weights for quitters	nWTS525v	nWTS565v
Rescaled waves 3–4 longitudinal weights for all tobacco users	nWTS625v	nWTS665v
Rescaled waves 3–4 longitudinal weights for all respondents	nWTS125v	nWTS165v

⋆ The dedicated Australian CCV sample was not recontacted at wave 4, and there was thus no need to compute separate weights for Australian NVP users as was done at waves 1–3.
† Too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 4. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 1–4 longitudinal weights for those user groups. To prevent confusion and misinterpretation, it was decided to also no longer compute waves 1–4 longitudinal weights for all tobacco users, as these would consists of different individuals than at other waves.
‡ Too few Canadian, British and Australian exclusive NVP users recruited at waves 1 or 2 were retained and interviewed at wave 4, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–4 longitudinal weights those individuals.

Table 14: Longitudinal sampling weights for wave 4 of the 4CV Survey

1. The first 7 sets of cross-sectional weights (see section 4.1.2 and column 2 of table 13) were computed for the main sample. That sample consists of 7991 respondents (2107 from the US, 2120 from Canada, 2267 from England and 1497 from Australia; see table 15). It includes all respondents except 210 long-term quitters (quit more than 5 years ago: 77 from the US, 64 from Canada, 45 from England and 24 from Australia) who are not NVP/HTP users, and the 71 respondents (21 from the US, 24 from Canada, 21 from England, and 5 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).

2. The next 7 sets of cross-sectional weights (see section 4.1.3 and column 3 of table 13) were computed for the reduced US sample. This sample is the main US sample of 2107 respondents (mentioned above) minus the 30 NVP users recruited via Ipsos' non-probability Web panel at wave 1 and successfully recontacted at waves 2, 3 and 4; hence, the reduced US sample consists of 2077 respondents. Those 30 respondents are the only US respondents not recruited using a probability based method, and those sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired. Since this only applies to US respondents, no such weights were computed for respondents from Canada, England or Australia.

4.1.1 User groups

As in previous waves, a key step in the computation of all sets of wave 4 cross-sectional weights consisted in dividing respondents into 4 user groups (variable `nUserGroup` in the dataset), and then further dividing quitters into 4 sub-groups (variable `nQuitGroup` in the dataset). However, how those groups and sub-groups are defined was slightly updated at wave 4. This was done to include (at least monthly) users of snus and of tobacco-free oral nicotine products products (ONPs) in Canada, the US and England; no change in Australia.

Hence, starting at wave 4, the user group definitions of section 3.1.1 were updated to the following for Canada, the US and England:

- To be classified as a cigarette only smoker (group *i*), a respondent had to (at the time of data collection) smoke cigarettes (at least occasionally), smoked at least 100 cigarettes in their lifetime, use NVP/HTP (this includes e-cigarettes) and/or snus and/or ONPs less than monthly or not at all.

  Note that the comment made in section 1.1.1 about including those less than monthly but occasional smokers also applies here, and below.

- To be classified as a dual user (group *ii*), a respondent had to (at the time of data collection) smoke cigarettes (at least occasionally), smoked at least 100 cigarettes in their lifetime, and use NVP/HTP (this includes e-cigarettes) and/or snus and/or ONPs at least monthly.

  Note that the comment made in section 1.1.1 about the use of the at "least monthly" criteria for the use of NVP being a limitation of our study also applies here, and below.

- To be classified as an exclusive NVP/HTP user (group *iii*), a respondent had to (at the time of data collection) use NVP/HTP (this includes e-cigarettes) and/or snus and/or ONPs at least monthly, and satisfy one or more of the following:
  - never having been a smoker
  - smoked less than 100 cigarettes in their lifetime
  - quit smoking cigarettes more than 5 years ago

- To be classified as a quitter (group *iv*), a respondent had to (at the time of data collection) have quit smoking cigarettes within the last 5 years, and smoked at least 100 cigarettes in their lifetime prior to that. Those who were using NVP/HTP (this includes e-cigarettes) and/or snus and/or ONPs at least monthly were then further sub-classified in either group *iv.a* or *iv.b* depending on how long ago they had quit smoking; whereas those who were using NVP/HTP and/or snus and/or ONPs less than monthly or not at all were then further sub-classified in either group *iv.c* or *iv.d* depending on how long ago they had quit smoking.

The numbers of respondents in each of those user groups are given in table 15.

User group†	US		Canada	England	Australia
	Main‡	Reduced‡			
Cigarette smokers					
Cigarette only	1134	1116	951	804	889
Dual users	301	297	673	1090	294
Total	1435	1413	1624	1894	1183
Exclusive NVP/HTP users	200	196	68	84	37
Quitters					
NVP/HTP users	189	188	140	140	125
Non users	283	280	288	149	152
Total	472	468	428	289	277
Total	2107	2077	2120	2267	1497

† Variables `nUserGroup` and `nQuitGroup` in the dataset
‡ See notes below table 13 and beginning of section 4.1

Table 15: Wave 4 respondents by country and user group.

4.1.2 Cross-sectional weights for the main sample

1- Variable `nWTS100v` contains the wave 4 cross-sectional inflation weights for the main sample of 7991 respondents (2107 from the US, 2120 from Canada, 2267 from England and 1497 from Australia; see table 15). How those weights are computed changed a little at wave 4 compare to the three previous waves (sections 1.1.2, 2.1.1 and 3.1.2).

As in previous waves, respondents were still divided into 4 user groups (i.e., groups *i*–*iv* of section 4.1.1), and quitters were further subdivided into 4 subgroups (i.e., *iv.a*–*iv.d*). Those groups were then crossed with gender, age, geographic region, education, ethnicity (US only) and language (Canada only); yielding the following cross-tabs: user group × gender, user group × age group, user group × geographic region, user group × education, user group × ethnicity (US only) and user group × language (Canada only). Benchmark figures were then obtained for those cross-tabs and for the four quitter subgroups. In the US, the 2022 National Health Interview Survey (NHIS) was used; whereas the 2021 Canadian Community Health Survey (CCHS) and the 2019 National Drug Strategy Household Survey (NDSHS) were used for Canada and Australia. In England, the 2021 & 2022 Opinions and Lifestyle Survey (OLS) were combined with waves 190–194 (Feb–Jun 2022)

of the Smoking Toolkit Study (STS) to obtain the calibration figures. These figures are given in appendix A.2. Note that, for Canada, the CCHS was used instead of CTADS and CTNS (which were used at prior waves), as those surveys did not contain all the variable we wanted to include in our weight calculations (e.g., CCHS includes education, whereas the other surveys do not). The raking algorithm (appendix A.1) was then used to calibrate the weights using the above mentioned cross-tabs and the four quitter subgroups; this was done separately for each country.

The previously described weight calculation procedure is the same as the one used at all previous three waves. The weakness of this procedure is that all respondents are assigned the same initial weight $w_i^{(0)} = 1$ at the beginning of the raking algorithm. Assuming that all respondents share the same initial weight, and thus the same initial selection probability is obviously incorrect. Fortunately, the raking algorithm mostly corrects this. Furthermore, until recently, there was little in terms of proper statistical methods to compute and assign initial weights to respondents of a non-probability based panel, like the ones used to obtain our Canadian, English and Australian samples. However, in their 2020 paper, Chen, Li & Wu introduced a method for computing sampling weights for non-probability survey. It was thus decided to use their new method to compute initial weights, and then use those weights as the $w_i^{(0)}$ starting weights for our raking algorithm of appendix A.1.

As detailed in appendix A.4, computation of those initial weights essentially consists in fitting a logistic regression model where the outcome is the probability of selection for individuals in our ITC sample as a function of various covariates. Those covariates are either: 1) variables used in the sampling design (e.g., user group in our case), 2) variables that impact the probability that a given respondent will respond (e.g., sex, age, education, race, socio economic status, etc.), or 3) or a combination of those two things. In our case, this corresponds to the covariates we were already using in our raking algorithm (see the various crosstabs mentioned above). Note that for age we used the respondents actual ages (i.e., continuous), as opposed to age groups for the raking algorithm. Compared to a traditional logistic regression model, the one used by Chen, Li & Wu is based on the joint dataset comprising of the surveys for which we desire to compute the initial weights (i.e., our ITC survey in the present case) and a reference (ideally) probability-based survey (i.e., the various national surveys mentioned above in the present case). That reference survey must include sampling weights as those will also be incorporated in the regression model. Furthermore, because the logistic model is based on the joint dataset, it is essential that the covariates/questions used in that model are the same in both surveys (likewise for the response options). Though this is an obvious limitation, it was not an issue in our case.

In summary, in Canada, the US and Australia, initial weights were computed using the method of Chen, Li & Wu (2020), as described in the above paragraph and in appendix A.4. The initial weights were then used as the $w_i^{(0)}$ starting weights of our raking algorithm. Except for this change and the fact that updated benchmark figures on smoking and tobacco use (see paragraph above) were used, calculations of the wave 4 cross-sectional inflation weights essentially proceeded as in prior waves. This was very much intentional, as our aim was to both improve on our weights calculations, but to also ensure consistency over time thus ensuring that our data can be combined and compared over waves. As mentioned in the above paragraph, the method of Chen, Li & Wu (2020) requires both surveys/datasets: our ITC survey and the national benchmark survey; that is, the actual dataset, as opposed to just the benchmark figures required for the raking algorithm and give in A.2. Unfortunately, it was not possible for us to gain access to the Opinions and Lifestyle Survey (OLS). Consequently, computation of the English wave 4 cross-sectional inflation weights proceeded the same way as in prior waves; with the obvious exception that updated benchmark figures were used.

2- Variable nWTS201v contains the rescaled wave 4 cross-sectional weights for the 6136 (1435 from the US, 1624 from Canada, 1894 from England and 1183 from Australia; see table 15) respondents who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 4 data collection. Starting at wave 4, some of those cigarette smokers were also using snus and/or ONPs. These are simply the wave 4 cross-sectional inflation weights (variable nWTS100v) of those 6136 respondents rescaled to sum to sample size in each country (i.e., 1435 in the US, 1624 in Canada, 1894 in England and 1183 in Australia). These weights are designed to make these 1624 Canadian cigarette smokers representative of the Canadian population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 4 data collection; likewise for the US, England and Australia.

3- Variable nWTS301v contains the rescaled wave 4 cross-sectional weights for the 3341 (690 from the US, 881 from Canada, 1314 from England and 456 from Australia; see table 15) respondents who were at least monthly NVP/HTP users at the time of wave 4 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs. These are simply the wave 4 cross-sectional inflation weights (variable nWTS100v) of those 3341 respondents rescaled to sum to sample size in each country. These weights are designed to make these 881 Canadian NVP/HTP users representative of the Canadian population of at least monthly NVP/HTP users at the time of wave 4 data collection; likewise for the US, England and Australia.

4- Variable nWTS401v contains the rescaled wave 4 cross-sectional weights for the 2358 (301 from the US, 673 from Canada, 1090 from England and 294 from Australia; see table 15) respondents who were dual users at the time of wave 4 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs. These are simply the wave 4 cross-sectional inflation weights (variable nWTS100v) of those 2358 respondents rescaled to sum to sample size in each country. These weights are designed to make these 673 Canadian dual users representative of the Canadian population of dual users at the time of wave 4 data collection; likewise for the US, England and Australia.

5- Variable nWTS501v contains the rescaled wave 4 cross-sectional weights for the 1466 (472 from the US, 428 from Canada, 289 from England and 277 from Australia; see table 15) quitters (within the last 5 years) at the time of wave 4 data collection. Starting at wave 4, some of those quitters were also using snus and/or ONPs. These are simply the wave 4 cross-sectional inflation weights (variable nWTS100v) of those 1466 respondents rescaled to sum to sample size in each country. The cautionary notes of variable kWTS501v and of section 6.3 still apply here.

6- Variable nWTS601v contains the rescaled wave 4 cross-sectional weights for the 7119 (1824 from the US, 1832 from Canada, 2118 from England and 1345 from Australia; see table 15) respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users[15]) at the time of wave 4 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs. These are simply the wave 4 cross-sectional inflation weights (variable nWTS100v) of those 7119 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1832 Canadian tobacco users representative of the Canadian population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users) at the time of wave 4 data collection; likewise for the US, England and Australia.

7- Variable nWTS101v contains the rescaled wave 4 cross-sectional weights for the main sample of 7991 respondents (2107 from the US, 2120 from Canada, 2267 from England and 1497 from Australia; see

[15]Note that some of those NVP/HTP users are individuals who have quit smoking traditional cigarette.

table 15). These are simply the wave 4 cross-sectional inflation weights (variable nWTS100v) of those 7991 respondents rescaled to sum to sample size in each country (i.e., 2107 in the US, 2120 in Canada, 2267 in England and 1497 in Australia). These weights are designed to make these 2120 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters at the time of wave 4 data collection; likewise for the US, England and Australia.

As mentioned in the description of variable kWTS101v, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the nWTS101v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (e.g., linear regression and logistic regression) from that same joint population is probably much more questionable.

4.1.3 Cross-sectional weights for the reduced US sample

As mentioned at the beginning of section 4.1, the following 7 sets of weights only apply to the US sample.

1- Variable nWTS102v contains the wave 4 cross-sectional inflation weights for the reduced US sample of 2077 respondents. As noted below table 13, the reduced sample is the main sample of 2107 respondents minus the 30 NVP users recruited by Ipsos at wave 1 (see section 1.1.3), and retained at waves 2, 3 and 4. Those NVP users are identifiable via the variable nOWNERID[16] in the dataset. Those at least monthly NVP users were recruited via a non-probability based panel, and the 7 sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired for their specific analyses.

   These weights were computed the same way as the wave 4 cross-sectional inflation weights for the main sample (see variable nWTS100v in section 4.1.2).

2- Variable nWTS203v contains the rescaled wave 4 cross-sectional weights for the 1413 respondents (see table 15) from the reduced US sample who were cigarette smokers at the time of wave 4 data collection. Starting at wave 4, some of those cigarette smokers were also using snus and/or ONPs. These are simply the wave 4 cross-sectional inflation weights (variable nWTS102v) of those respondents rescaled to sum to sample size (i.e, 1413). These weights are designed to make these 1413 US cigarette smokers representative of the US population of cigarette smokers at the time of wave 4 data collection. In other words, variable nWTS203v is the analogue of variable nWTS201v, but for the reduced sample.

3- Variable nWTS303v is the analogue of variable nWTS301v, but for the reduced US sample of 681 respondents who were at least monthly NVP/HTP users at the time of wave 4 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs.

4- Variable nWTS403v is the analogue of variable nWTS401v, but for the reduced US sample of 297 respondents who were dual users at the time of wave 4 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs.

5- Variable nWTS503v is the analogue of variable nWTS501v, but for the reduced US sample of 468 respondents who were quitters at the time of wave 4 data collection. Starting at wave 4, some of those quitters were also using snus and/or ONPs.

[16] Use nOWNERID!=05 to exclude the 30 US NVP users recruited by Ipsos, where != means not equal to.

6- Variable `nWTS603v` is the analogue of variable `nWTS601v`, but for the reduced US sample of 1797 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users) at the time of wave 4 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs.

7- Variable `nWTS103v` is the analogue of variable `nWTS101v`, but for the reduced US sample of 2077 respondents. Hence, the cautionary notes listed for variable `nWTS101v` also apply here.

4.2 Waves 1–4 longitudinal sampling weights

As mentioned at the beginning of section 4 (table 14), 30 sets of longitudinal weights were computed at wave 4 of the 4CV Survey. Six of those (see top third of table 14) were computed for respondents recruited at wave 1, and then successfully retained and interviewed at waves 2, 3 and 4. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 1–4 longitudinal weights are designed to make the sample representative at the time of wave 1. If the aim is to ensure that the sample is representative at the time of wave 4 data collection, then the wave 4 cross-sectional sampling weights computed in section 4.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

As mentioned in table 14, too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 4, and it was decided to no longer compute waves 1–4 longitudinal weights for those user groups.

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette only	321	321	394	393	262
Dual users	118	108	127	121	41
Total	439	429	521	514	303

† Variables `kUserGroup` in the dataset

Table 16: Wave 1 respondents successfully recontacted at wave 4 by user group.

4.2.1 Waves 1–4 longitudinal weights for the main sample

Waves 1–4 longitudinal inflation weights were computed for the 1777 respondents (439 from the US, 521 from Canada, 514 from England and 303 from Australia) from the wave 1 main sample that were successfully recontacted at waves 2, 3 and 4. These inflation weights are the basis of the 3 set of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 1–4 longitudinal inflation weights are the waves 1–3 longitudinal inflation weights (computed section 3.2.1) adjusted for attrition between waves 3 and 4. As with the waves 1–3 longitudinal

inflation weights, computation of the waves 1–4 longitudinal inflation weights followed the same steps as described in section 1.1.2. Since these weights are meant to make these 1777 respondents representative of the population at the time of wave 1 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 1 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 4 cross-sectional weights for the main sample (see section 4.1.2). Note that respondents were divided into the 4 user groups based on their wave 1 status (variables `kUserGroup`), and that age at wave 1 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Though computation followed the same steps as described in section 1.1.2, too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 4, and it was decided to no longer compute waves 1–4 longitudinal weights for those user groups.

1- Variable `nWTS221v` contains the rescaled waves 1–4 longitudinal weights for the 1777 respondents (439 from the US, 521 from Canada, 514 from England and 303 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 1 data collection and successfully recontacted at wave 4. These are simply the waves 1–4 longitudinal inflation weights (computed above) of those 1777 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 1 data collection.

2- Variable `nWTS421v` contains the rescaled waves 1–4 longitudinal weights for the 407 respondents (118 from the US, 127 from Canada, 121 from England and 41 from Australia) who were dual users at the time of wave 1 data collection and successfully recontacted at wave 4. These are simply the waves 1–4 longitudinal inflation weights (computed above) of those 407 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 1 data collection.

3- Variable `nWTS121v` contains the rescaled waves 1–4 longitudinal weights for all 1777 respondents (439 from the US, 521 from Canada, 514 from England and 303 from Australia) from the wave 1 main sample that were successfully recontacted at wave 4. As mentioned at the top of section 4.2 and in table 14, too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 4, and it was decided to no longer compute waves 1–4 longitudinal weights for those user groups. Consequently, variable `nWTS121v` is the same as `nWTS221v`.

4.2.2 Waves 1–4 longitudinal weights for the US reduced sample

The following 3 sets of waves 1–4 longitudinal weights were computed for the 429 respondents from the wave 1 US reduced sample (see third column of table 16) that were successfully recontacted at waves 2, 3 and 4. Computation followed the same steps as described in section 4.2.1, and the following weights are also designed to make those 429 respondents representative at the time of wave 1.

As in section 4.2.1, no weights were computed for exclusive NVP users and quitters recruited at wave 1 and retained/interviewed at wave 4, as there were simply too few of them to be representative.

1- Variable `nWTS261v` is the analogue of variable `nWTS221v`, but for the reduced US sample of 429 respondents who were cigarette smokers at the time of wave 1 data collection.

2- Variable `nWTS461v` is the analogue of variable `nWTS421v`, but for the reduced US sample of 108 respondents who were dual users at the time of wave 1 data collection.

3- Variable `nWTS161v` is the analogue of variable `nWTS121v`, but for the reduced US sample of 429 respondents. In the present case, it is the same as variable `nWTS261v`, since no waves 1–4 longitudinal weights were computed for exclusive NVP users and quitters.

4.3 Waves 2–4 longitudinal sampling weights

As mentioned at the beginning of section 4, 30 sets of longitudinal weights were computed at wave 4 of the 4CV Survey. Twelve of those (see middle third of table 14) were computed for respondents interviewed at wave 2, and then successfully retained and interviewed at both wave 3 and wave 4. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 2—4 longitudinal weights are designed to make the sample representative at the time of wave 2. If the aim is to ensure that the sample is representative at the time of wave 4 data collection, then the wave 4 cross-sectional sampling weights computed in section 4.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

As mentioned in table 14, too few Canadian, British and Australian exclusive NVP users recruited at waves 1 or 2 were retained and interviewed at wave 4, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–4 longitudinal weights those individuals.

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette only	526	511	520	416	338
Dual users	168	160	180	165	65
Exclusive NVP users	30	27	0	0	0
Quitters	184	178	166	0	80
Total	908	876	866	581	483

† Variables `lUserGroup` and `lQuitGroup` in the dataset

Table 17: Wave 2 respondents successfully recontacted at wave 4 by user group.

4.3.1 Waves 2–4 longitudinal weights for the main sample

Waves 2–4 longitudinal inflation weights were computed for the 2838 respondents (908 from the US, 866 from Canada, 581 from England and 483 from Australia) from the wave 2 main sample that were successfully recontacted at waves 3 and 4. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 2–4 longitudinal inflation weights are the waves 2–3 longitudinal inflation weights (computed in section 3.3.1) adjusted for attrition between waves 3 and 4. Consequently, computation of those

longitudinal inflation weights followed the same steps as described in section 2.1.1. Since these weights are meant to make these 2838 respondents representative of the population at the time of wave 2 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 2 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 4 cross-sectional weights for the main sample (see section 4.1.2). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 2 status (variables `lUserGroup` and `lQuitGroup`), and that age at wave 2 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Though computation followed the same steps as described in section 2.1.1, too few Canadian, British and Australian exclusive NVP users from the wave 2 main sample were retained and interviewed at wave 4, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–4 longitudinal weights those individuals.

1- Variable `nWTS223v` contains the rescaled waves 2–4 longitudinal weights for the 2378 respondents (694 from the US, 700 from Canada, 581 from England and 403 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 2 data collection and successfully recontacted at waves 3 and 4. These are simply the waves 2–4 longitudinal inflation weights (computed above) of those 2378 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 2 data collection.

2- Variable `nWTS323v` contains the rescaled waves 2–4 longitudinal weights for the 774 respondents (289 from the US, 226 from Canada, 165 from England and 94 from Australia) who were at least monthly NVP users at the time of wave 2 data collection and successfully recontacted at waves 3 and 4. These are simply the waves 2–4 longitudinal inflation weights (computed above) of those 774 respondents rescaled to sum to sample size in each country.

   Cautionary note: As mentioned at the beginning of section 4.3 and in table 17, too few Canadian, British and Australian exclusive NVP users recruited at waves 1 or 2 were retained and interviewed at wave 4, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–4 longitudinal weights those individuals (thus no `nWTS323v` weights were computed for those individuals). This is of key importance, as this means that the target population for the `nWTS323v` weights is not the same for all 4 countries. That is, these weights are designed to make American NVP users representative of the population of at least monthly NVP users. For Canada and Australia, these weights are designed to make NVP users representative of the population consisting of dual users and quitters who use NVPs at least monthly. For England, these weights are designed to make NVP users representative of the population consisting of dual users.

3- Variable `nWTS423v` contains the rescaled waves 2–4 longitudinal weights for the 578 respondents (168 from the US, 180 from Canada, 165 from England and 65 from Australia) who were dual users at the time of wave 2 data collection and successfully recontacted at waves 3 and 4. These are simply the waves 2–4 longitudinal inflation weights (computed above) of those 578 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 2 data collection.

4- Variable `nWTS523v` contains the rescaled waves 2–4 longitudinal weights for the 430 respondents (184 from the US, 166 from Canada, 0 from England and 80 from Australia) who had quit smoking

cigarettes at the time of wave 2 data collection and successfully recontacted at waves 3 and 4. These are simply the waves 2–4 longitudinal inflation weights (computed above) of those 430 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country (with the exception of England obviously[17]) at the time of wave 2 data collection. The cautionary notes of variable kWTS501v and of section 6.3 still apply here.

5- Variable nWTS623v contains the rescaled waves 2–4 longitudinal weights for the 2574 respondents (815 from the US, 746 from Canada, 581 from England and 432 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection and successfully recontacted at waves 3 and 4. These are simply the waves 2–4 longitudinal inflation weights (computed above) of those 2574 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 2 data collection.

   The cautionary note about exclusive NVP users and English quitters stated with variable nWTS323v also applies here.

6- Variable nWTS123v contains the rescaled waves 2–4 longitudinal weights for all 2838 respondents (908 from the US, 866 from Canada, 581 from England and 483 from Australia) from the wave 2 main sample that were successfully recontacted at waves 3 and 4. These are simply the waves 2–4 longitudinal inflation weights (computed above) of those 2838 respondents rescaled to sum to sample size in each country. These weights are designed to make these 866 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 2 data collection; likewise for the US and Australia (within the last 2 years). For England, these weights are designed to make these 581 British tobacco users representative of the British population of cigarette smokers and dual users.

   The cautionary note about exclusive NVP users and English quitters stated with variable nWTS323v also applies here.

   As mentioned in the description of variable nWTS101v, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the nTS123v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

4.3.2 Waves 2–4 longitudinal weights for the US reduced sample

The following 6 sets of waves 2–4 longitudinal weights were computed for the 876 respondents from the wave 2 US reduced sample (see third column of table 17) that were successfully recontacted at waves 3 and 4. Computation followed the same steps as described in section 4.3.1, and the following weights are also designed to make those 876 respondents representative at the time of wave 2.

1- Variable nWTS263v is the analogue of variable nTS223v, but for the reduced US sample of 671 respondents who were cigarette smokers at the time of wave 2 data collection.

[17] Recall that too few British quitter recruited at waves 1 or 2 were retained and interviewed at wave 4. Consequently, it was decided to no longer compute waves 2–4 longitudinal for those individuals.

2- Variable `nWTS363v` is the analogue of variable `nWTS323v`, but for the reduced US sample of 277 respondents who were at least monthly NVP users at the time of wave 2 data collection.

3- Variable `nWTS463v` is the analogue of variable `nWTS423v`, but for the reduced US sample of 160 respondents who were dual users at the time of wave 2 data collection.

4- Variable `nWTS563v` is the analogue of variable `nWTS523v`, but for the reduced US sample of 178 respondents who were quitters at the time of wave 2 data collection.

5- Variable `nWTS663v` is the analogue of variable `nWTS623v`, but for the reduced US sample of 788 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection.

6- Variable `nWTS163v` is the analogue of variable `nWTS123v`, but for the reduced US sample of 876 respondents.

4.4 Waves 3–4 longitudinal sampling weights

As mentioned at the beginning of section 4, 30 sets of longitudinal weights were computed at wave 4 of the 4CV Survey. Twelve of those (see bottom third of table 14) were computed for respondents interviewed at wave 3, and then successfully retained and interviewed at wave 4. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 3—4 longitudinal weights are designed to make the sample representative at the time of wave 3. If the aim is to ensure that the sample is representative at the time of wave 4 data collection, then the wave 4 cross-sectional sampling weights computed in section 4.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette only	703	686	826	589	520
Dual users	170	164	346	248	96
Exclusive NVP users	136	133	38	49	24
Quitters	247	242	233	161	123
Total	1256	1225	1443	1047	763

† Variables `mUserGroup` and `mQuitGroup` in the dataset

Table 18: Wave 3 respondents successfully recontacted at wave 4 by user group.

4.4.1 Waves 3–4 longitudinal weights for the main sample

Waves 3–4 longitudinal inflation weights were computed for the 4509 respondents (1256 from the US, 1443 from Canada, 1047 from England and 763 from Australia) from the wave 3 main sample that were successfully recontacted at wave 4. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal

analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 3–4 longitudinal inflation weights are the wave 3 cross-sectional inflation weights (i.e., variable mWTS100v, computed in section 3.1.2) adjusted for attrition between waves 3 and 4. Consequently, computation of those longitudinal inflation weights followed the same steps as described in section 3.1.2. Since these weights are meant to make these 4509 respondents representative of the population at the time of wave 3 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 3 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 4 cross-sectional weights for the main sample (see section 4.1.2). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 3 status (variables `mUserGroup` and `mQuitGroup`), and that age at wave 3 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

1- Variable `nWTS225v` contains the rescaled waves 3–4 longitudinal weights for the 3498 respondents (873 from the US, 1172 from Canada, 837 from England and 616 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 3 data collection and successfully recontacted at wave 4. These are simply the waves 3–4 longitudinal inflation weights (computed above) of those 3498 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 3 data collection.

2- Variable `nWTS325v` contains the rescaled waves 3–4 longitudinal weights for the 1423 respondents (459 from the US, 440 from Canada, 364 from England and 160 from Australia) who were at least monthly NVP users at the time of wave 3 data collection and successfully recontacted at wave 4. These are simply the waves 3–4 longitudinal inflation weights (computed above) of those 1423 respondents rescaled to sum to sample size in each country. These weights are designed to make these NVP users representative of the population of at least monthly NVP users of their respective country at the time of wave 3 data collection.

3- Variable `nWTS425v` contains the rescaled waves 3–4 longitudinal weights for the 860 respondents (170 from the US, 346 from Canada, 248 from England and 96 from Australia) who were dual users at the time of wave 3 data collection and successfully recontacted at wave 4. These are simply the waves 3–4 longitudinal inflation weights (computed above) of those 860 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 3 data collection.

4- Variable `nWTS525v` contains the rescaled waves 3–4 longitudinal weights for the 764 respondents (247 from the US, 233 from Canada, 161 from England and 123 from Australia) who had quit smoking cigarettes at the time of wave 3 data collection and successfully recontacted at wave 4. These are simply the waves 3–4 longitudinal inflation weights (computed above) of those 764 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country at the time of wave 3 data collection. The cautionary notes of variable kWTS501v and of section 6.3 still apply here.

5- Variable nWTS625v contains the rescaled waves 3–4 longitudinal weights for the 4061 respondents (1162 from the US, 1266 from Canada, 953 from England and 680 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 3 data collection and successfully recontacted at wave 4. These are simply the waves 3–4 longitudinal inflation weights (computed above) of those 4061 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 3 data collection.

6- Variable nWTS125v contains the rescaled waves 3–4 longitudinal weights for all 4509 respondents (1256 from the US, 1443 from Canada, 1047 from England and 763 from Australia) from the wave 3 main sample that were successfully recontacted at wave 4. These are simply the waves 3–4 longitudinal inflation weights (computed above) of those 4509 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1443 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 3 data collection; likewise for the US, England and Australia.

   As mentioned in the description of variable nWTS101v, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the nWTS125v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

4.4.2 Waves 3–4 longitudinal weights for the US reduced sample

The following 6 sets of waves 3–4 longitudinal weights were computed for the 1225 respondents from the wave 3 US reduced sample (see third column of table 18) that were successfully recontacted at wave 4. Computation followed the same steps as described in section 4.4.1, and the following weights are also designed to make those 1225 respondents representative at the time of wave 3.

1- Variable nWTS265v is the analogue of variable nTS225v, but for the reduced US sample of 850 respondents who were cigarette smokers at the time of wave 3 data collection.

2- Variable nWTS365v is the analogue of variable nWTS325v, but for the reduced US sample of 448 respondents who were at least monthly NVP users at the time of wave 3 data collection.

3- Variable nWTS465v is the analogue of variable nWTS425v, but for the reduced US sample of 164 respondents who were dual users at the time of wave 3 data collection.

4- Variable nWTS565v is the analogue of variable nWTS525v, but for the reduced US sample of 242 respondents who were quitters at the time of wave 3 data collection.

5- Variable nWTS665v is the analogue of variable nWTS625v, but for the reduced US sample of 1134 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 3 data collection.

6- Variable nWTS165v is the analogue of variable nWTS125v, but for the reduced US sample of 1225 respondents.

5 Wave 5 weights

Fourteen sets of cross-sectional weights and 42 sets of longitudinal weights were computed at wave 5 of the 4CV Survey. These cross-sectional weights are listed in table 19, and their computation and how/when they should be used are detailed in section 5.1. Likewise, the longitudinal weights are listed in table 20, and their computation and use are detailed in sections 5.2 (waves 1–5 longitudinal weights), 5.3 (waves 2–5 longitudinal weights), 5.4 (waves 3–5 longitudinal weights) and 5.5 (waves 4–5 longitudinal weights).

All sampling weights for the 4CV Survey were computed using the statistical software R (r-project.org). As mentioned at the beginning of this document, these weights adjust for: 1) oversampling of 18–24 years old tobacco users (at waves 1–3 in US, Canada & England), 2) the oversampling of nicotine vaping product (NVP) users and e-cigarette users[18], 3) sample mis-representation, 4) non-response and 5) other biases. **It is thus essential to use weighted data, when performing any analyses using 4CV data**.

Weight	Variable Names	
	Main sample⋆	reduced US sample†
Wave 5 cross-sectional inflation weights	`oWTS100v`	`oWTS102v`
Rescaled wave 5 cross-sectional weights for cigarette smokers	`oWTS201v`	`oWTS203v`
Rescaled wave 5 cross-sectional weights for NVP users	`oWTS301v`	`oWTS303v`
Rescaled wave 5 cross-sectional weights for dual users	`oWTS401v`	`oWTS403v`
Rescaled wave 5 cross-sectional weights for quitters	`oWTS501v`	`oWTS503v`
Rescaled wave 5 cross-sectional weights for all tobacco users	`oWTS601v`	`oWTS603v`
Rescaled wave 5 cross-sectional weights for all respondents	`oWTS101v`	`oWTS103v`

⋆ The main sample consists of 8219 respondents (2094 from the US, 2369 from Canada, 2275 from England and 1481 from Australia). It includes all respondents except 159 long-term quitters (quit more than 5 years ago: 75 from the US, 58 from Canada, 12 from England and 14 from Australia) who are not NVP/HTP users, and the 71 respondents (25 from the US, 27 from Canada, 12 from England, and 7 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).
† The reduced sample is the main US sample of 2094 respondents minus the 18 NVP users recruited by Ipsos at wave 1 and successfully recontacted at wave 5; hence, the reduced US sample consists of 2076 respondents.

Table 19: Cross-sectional sampling weights for wave 5 of the 4CV Survey

5.1 Cross-sectional sampling weights

Similarly to previous waves, the 14 sets of cross-sectional weights can be divided into 2 categories:

1. The first 7 sets of cross-sectional weights (see section 5.1.1 and column 2 of table 19) were computed for the main sample. That sample consists of 8219 respondents (2094 from the US, 2369 from Canada, 2275 from England and 1481 from Australia; see table 21). It includes all respondents

[18] At all waves of US, Canada & England, and waves 1–3 of Australia

Weight	Variable Names: Main sample	Variable Names: reduced US sample
Rescaled waves 1–5 longitudinal weights for cigarette smokers	oWTS221v	oWTS261v
Rescaled waves 1–5 longitudinal weights for NVP users	n/a†	n/a†
Rescaled waves 1–5 longitudinal weights for dual users	oWTS421v	oWTS461v
Rescaled waves 1–5 longitudinal weights for quitters	n/a†	n/a†
Rescaled waves 1–5 longitudinal weights for all tobacco users	n/a†	n/a†
Rescaled waves 1–5 longitudinal weights for all respondents	oWTS121v	oWTS161v
Rescaled waves 2–5 longitudinal weights for cigarette smokers	oWTS223v	oWTS263v
Rescaled waves 2–5 longitudinal weights for NVP users	oWTS323v‡	oWTS363v
Rescaled waves 2–5 longitudinal weights for dual users	oWTS423v	oWTS463v
Rescaled waves 2–5 longitudinal weights for quitters	oWTS523v‡	oWTS563v
Rescaled waves 2–5 longitudinal weights for all tobacco users	oWTS623v‡	oWTS663v
Rescaled waves 2–5 longitudinal weights for all respondents	oWTS123v	oWTS163v
Rescaled waves 3–5 longitudinal weights for cigarette smokers	oWTS225v	oWTS265v
Rescaled waves 3–5 longitudinal weights for NVP users	oWTS325v	oWTS365v
Rescaled waves 3–5 longitudinal weights for dual users	oWTS425v	oWTS465v
Rescaled waves 3–5 longitudinal weights for quitters	oWTS525v	oWTS565v
Rescaled waves 3–5 longitudinal weights for all tobacco users	oWTS625v	oWTS665v
Rescaled waves 3–5 longitudinal weights for all respondents	oWTS125v	oWTS165v
Rescaled waves 4–5 longitudinal weights for cigarette smokers	oWTS227v	oWTS267v
Rescaled waves 4–5 longitudinal weights for NVP users	oWTS327v	oWTS367v
Rescaled waves 4–5 longitudinal weights for dual users	oWTS427v	oWTS467v
Rescaled waves 4–5 longitudinal weights for quitters	oWTS527v	oWTS567v
Rescaled waves 4–5 longitudinal weights for all tobacco users	oWTS627v	oWTS667v
Rescaled waves 4–5 longitudinal weights for all respondents	oWTS127v	oWTS167v

⋆ The dedicated Australian CCV sample was not recontacted at waves 4 and 5, and there was thus no need to compute separate weights for Australian NVP users as was done at waves 1–3.
† Too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 5. Since the representativeness of such small samples is highly questionnable, it was decided not to compute waves 1–5 longitudinal weights for those user groups. To prevent confusion and misinterpretation, it was decided to also not compute waves 1–5 longitudinal weights for all tobacco users, as these would consists of different individuals than at other waves.
‡ Too few Canadian, British and Australian exclusive NVP users recruited at waves 1 or 2 were retained and interviewed at wave 5, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided not to compute waves 2–5 longitudinal weights those individuals.

Table 20: Longitudinal sampling weights for wave 5 of the 4CV Survey

except 159 long-term quitters (quit more than 5 years ago: 75 from the US, 58 from Canada, 12 from England and 14 from Australia) who are not NVP/HTP users, and the 71 respondents (25 from the US, 27 from Canada, 12 from England, and 7 from Australia) deemed to be fraudulent (also referred to as speeders in some 4CV documentation).

2. The next 7 sets of cross-sectional weights (see section 5.1.2 and column 3 of table 19) were computed for the reduced US sample. This sample is the main US sample of 2094 respondents (mentioned above) minus the 18 NVP users recruited via Ipsos' non-probability Web panel at wave 1 and successfully recontacted at waves 2, 3 and 4; hence, the reduced US sample consists of 2076 respondents. Those 18 respondents are the only US respondents not recruited using a probability based method, and those sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired. Since this only applies to US respondents, no such weights were computed for respondents from Canada, England or Australia.

The numbers of respondents in each of those user groups are given in table 21.

User group†	US		Canada	England	Australia
	Main‡	Reduced‡			
Cigarette smokers					
Cigarette only	1013	1003	925	773	731
Dual users	354	351	848	1140	378
Total	1367	1354	1773	1913	1109
Exclusive NVP/HTP users	216	213	158	106	81
Quitters					
NVP/HTP users	207	207	172	153	141
Non users	304	302	266	103	150
Total	511	509	438	256	291
Total	2094	2076	2369	2275	1481

† Variables `oUserGroup` and `oQuitGroup` in the dataset
‡ See notes below table 19 and beginning of section 5.1

Table 21: Wave 5 respondents by country and user group.

5.1.1 Cross-sectional weights for the main sample

1- Variable `oWTS100v` contains the wave 5 cross-sectional inflation weights for the main sample of 8219 respondents (2094 from the US, 2369 from Canada, 2275 from England and 1481 from Australia; see table 15). As discuseed in section 4.1.2, how those weights are computed changed a little at wave 4 compare to the three previous waves (sections 1.1.2, 2.1.1 and 3.1.2). With the exception that updated benchmark figures were used, computation of these cross-sectional inflation weights proceeded exactly as with the wave 4 cross-sectional inflation weights (section 4.1.2) and is described in the paragraphs below.

As in previous waves, respondents were still divided into 4 user groups (i.e., groups *i*–*iv* of section 4.1.1), and quitters were further subdivided into 4 subgroups (i.e., *iv.a*–*iv.d*). Those groups

were then crossed with gender, age, geographic region, education, ethnicity (US only) and language (Canada only); yielding the following cross-tabs: user group × gender, user group × age group, user group × geographic region, user group × education, user group × ethnicity (US only) and user group × language (Canada only). Benchmark figures were then obtained for those cross-tabs and for the four quitter subgroups. In the US, the 2022 & 2023 National Health Interview Survey (NHIS) was used; whereas the 2022 Canadian Community Health Survey (CCHS) and the 2022–23 National Drug Strategy Household Survey (NDSHS) were used for Canada and Australia. In England, the 2022 & 2023 Opinions and Lifestyle Survey (OLS) were combined with waves 213–217 (Jul–Nov 2024) of the Smoking Toolkit Study (STS) to obtain the calibration figures. These figures are given in appendix A.2. The raking algorithm (appendix A.1) was then used to calibrate the weights using the above mentioned cross-tabs and the four quitter subgroups; this was done separately for each country.

The previously described weight calculation procedure is the same as the one used at waves 1–3. The weakness of this procedure is that all respondents are assigned the same initial weight $w_i^{(0)} = 1$ at the beginning of the raking algorithm. Assuming that all respondents share the same initial weight, and thus the same initial selection probability is obviously incorrect. Fortunately, the raking algorithm mostly corrects this. Furthermore, until recently, there was little in terms of proper statistical methods to compute and assign initial weights to respondents of a non-probability based panel, like the ones used to obtain our Canadian, English and Australian samples. However, in their 2020 paper, Chen, Li & Wu introduced a method for computing sampling weights for non-probability survey. It was thus decided to use their new method to compute initial weights, and then use those weights as the $w_i^{(0)}$ starting weights for our raking algorithm of appendix A.1.

As detailed in appendix A.4, computation of those initial weights essentially consists in fitting a logistic regression model where the outcome is the probability of selection for individuals in our ITC sample as a function of various covariates. Those covariates are either: 1) variables used in the sampling design (e.g., user group in our case), 2) variables that impact the probability that a given respondent will respond (e.g., sex, age, education, race, socio economic status, etc.), or 3) or a combination of those two things. In our case, this corresponds to the covariates we were already using in our raking algorithm (see the various crosstabs mentioned above). Note that for age we used the respondents actual ages (i.e., continuous), as opposed to age groups for the raking algorithm. Compared to a traditional logistic regression model, the one used by Chen, Li & Wu is based on the joint dataset comprising of the surveys for which we desire to compute the initial weights (i.e., our ITC survey in the present case) and a reference (ideally) probability-based survey (i.e., the various national surveys mentioned above in the present case). That reference survey must include sampling weights as those will also be incorporated in the regression model. Furthermore, because the logistic model is based on the joint dataset, it is essential that the covariates/questions used in that model are the same in both surveys (likewise for the response options). Though this is an obvious limitation, it was not an issue in our case.

In summary, in Canada, the US and Australia, initial weights were computed using the method of Chen, Li & Wu (2020), as described in the above paragraph and in appendix A.4. The initial weights were then used as the $w_i^{(0)}$ starting weights of our raking algorithm. Except for this change and the fact that updated benchmark figures on smoking and tobacco use (see paragraph above) were used, calculations of the wave 5 cross-sectional inflation weights essentially proceeded as in waves 1–3 and exactly the same as in wave 4. This was very much intentional, as our aim was to both improve on our weights calculations, but to also ensure consistency over time thus ensuring that our data can be combined and compared over waves. As mentioned in the above paragraph, the method of Chen, Li &

Wu (2020) requires both surveys/datasets: our ITC survey and the national benchmark survey; that is, the actual dataset, as opposed to just the benchmark figures required for the raking algorithm and give in A.2. Unfortunately, it was not possible for us to gain access to the Opinions and Lifestyle Survey (OLS). Consequently, computation of the English wave 5 cross-sectional inflation weights proceeded the same way as in prior waves; with the obvious exception that updated benchmark figures were used.

2- Variable `oWTS201v` contains the rescaled wave 5 cross-sectional weights for the 6162 (1367 from the US, 1773 from Canada, 1913 from England and 1109 from Australia; see table 21) respondents who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 5 data collection. Starting at wave 4, some of those cigarette smokers were also using snus and/or ONPs. These are simply the wave 5 cross-sectional inflation weights (variable `oWTS100v`) of those 6162 respondents rescaled to sum to sample size in each country (i.e., 1367 in the US, 1773 in Canada, 1913 in England and 1109 in Australia). These weights are designed to make these 1773 Canadian cigarette smokers representative of the Canadian population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 5 data collection; likewise for the US, England and Australia.

3- Variable `oWTS301v` contains the rescaled wave 5 cross-sectional weights for the 3954 (777 from the US, 1178 from Canada, 1399 from England and 600 from Australia; see table 21) respondents who were at least monthly NVP/HTP users at the time of wave 5 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs. These are simply the wave 5 cross-sectional inflation weights (variable `oWTS100v`) of those 3954 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1178 Canadian NVP/HTP users representative of the Canadian population of at least monthly NVP/HTP users at the time of wave 5 data collection; likewise for the US, England and Australia.

4- Variable `oWTS401v` contains the rescaled wave 5 cross-sectional weights for the 2720 (354 from the US, 848 from Canada, 1140 from England and 378 from Australia; see table 21) respondents who were dual users at the time of wave 5 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs. These are simply the wave 5 cross-sectional inflation weights (variable `oWTS100v`) of those 2720 respondents rescaled to sum to sample size in each country. These weights are designed to make these 848 Canadian dual users representative of the Canadian population of dual users at the time of wave 5 data collection; likewise for the US, England and Australia.

5- Variable `oWTS501v` contains the rescaled wave 5 cross-sectional weights for the 1496 (511 from the US, 438 from Canada, 256 from England and 291 from Australia; see table 21) quitters (within the last 5 years) at the time of wave 5 data collection. Starting at wave 4, some of those quitters were also using snus and/or ONPs. These are simply the wave 5 cross-sectional inflation weights (variable `oWTS100v`) of those 1496 respondents rescaled to sum to sample size in each country. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

6- Variable `oWTS601v` contains the rescaled wave 5 cross-sectional weights for the 7396 (1790 from the US, 2103 from Canada, 2172 from England and 1331 from Australia; see table 21) respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users[19]) at the time of wave 5 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs. These are simply the wave 5 cross-sectional inflation weights (variable `oWTS100v`) of those

[19]Note that some of those NVP/HTP users are individuals who have quit smoking traditional cigarette.

7396 respondents rescaled to sum to sample size in each country. These weights are designed to make these 2103 Canadian tobacco users representative of the Canadian population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users) at the time of wave 5 data collection; likewise for the US, England and Australia.

7- Variable `oWTS101v` contains the rescaled wave 5 cross-sectional weights for the main sample of 8219 respondents (2094 from the US, 2369 from Canada, 2275 from England and 1481 from Australia; see table 21). These are simply the wave 5 cross-sectional inflation weights (variable `oWTS100v`) of those 8219 respondents rescaled to sum to sample size in each country (i.e., 2094 in the US, 2369 in Canada, 2275 in England and 1481 in Australia). These weights are designed to make these 2369 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters at the time of wave 5 data collection; likewise for the US, England and Australia.

   As mentioned in the description of variable `kWTS101v`, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the `oWTS101v` weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (e.g., linear regression and logistic regression) from that same joint population is probably much more questionable.

5.1.2 Cross-sectional weights for the reduced US sample

As mentioned at the beginning of section 5.1, the following 7 sets of weights only apply to the US sample.

1- Variable `oWTS102v` contains the wave 5 cross-sectional inflation weights for the reduced US sample of 2076 respondents. As noted below table 19, the reduced sample is the main sample of 2094 respondents minus the 18 NVP users recruited by Ipsos at wave 1 (see section 1.1.3), and retained at wave 5. Those NVP users are identifiable via the variable `nOWNERID`[20] in the dataset. Those at least monthly NVP users were recruited via a non-probability based panel, and the 7 sets of cross-sectional weights were thus computed to give data users the flexibility to remove those individuals if they so desired for their specific analyses.

   These weights were computed the same way as the wave 5 cross-sectional inflation weights for the main sample (see variable `oWTS100v` in section 5.1.1).

2- Variable `oWTS203v` contains the rescaled wave 5 cross-sectional weights for the 1354 respondents (see table 21) from the reduced US sample who were cigarette smokers at the time of wave 5 data collection. Starting at wave 4, some of those cigarette smokers were also using snus and/or ONPs. These are simply the wave 5 cross-sectional inflation weights (variable `oWTS102v`) of those respondents rescaled to sum to sample size (i.e, 1354). These weights are designed to make these 1354 US cigarette smokers representative of the US population of cigarette smokers at the time of wave 5 data collection. In other words, variable `oWTS203v` is the analogue of variable `oWTS201v`, but for the reduced sample.

3- Variable `oWTS303v` is the analogue of variable `oWTS301v`, but for the reduced US sample of 771 respondents who were at least monthly NVP/HTP users at the time of wave 5 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs.

[20] Use `nOWNERID`!=05 to exclude the 18 US NVP users recruited by Ipsos, where != means not equal to.

4- Variable oWTS403v is the analogue of variable oWTS401v, but for the reduced US sample of 351 respondents who were dual users at the time of wave 5 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs.

5- Variable oWTS503v is the analogue of variable oWTS501v, but for the reduced US sample of 509 respondents who were quitters at the time of wave 5 data collection. Starting at wave 4, some of those quitters were also using snus and/or ONPs.

6- Variable oWTS603v is the analogue of variable oWTS601v, but for the reduced US sample of 1774 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP/HTP users) at the time of wave 5 data collection. Starting at wave 4, some of those users were also using snus and/or ONPs.

7- Variable oWTS103v is the analogue of variable oWTS101v, but for the reduced US sample of 2076 respondents. Hence, the cautionary notes listed for variable oWTS101v also apply here.

5.2 Waves 1–5 longitudinal sampling weights

As mentioned at the beginning of section 5 (table 20), 42 sets of longitudinal weights were computed at wave 5 of the 4CV Survey. Six of those (see top quarter of table 20) were computed for respondents recruited at wave 1, and then successfully retained and interviewed at waves 2, 3, 4 and 5. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 1–5 longitudinal weights are designed to make the sample representative at the time of wave 1. If the aim is to ensure that the sample is representative at the time of wave 5 data collection, then the wave 5 cross-sectional sampling weights computed in section 5.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

As mentioned in table 20, too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 5, and it was decided to no longer compute waves 1–5 longitudinal weights for those user groups.

5.2.1 Waves 1–5 longitudinal weights for the main sample

Waves 1–5 longitudinal inflation weights were computed for the 4509 respondents (1256 from the US, 1443 from Canada, 1047 from England and 763 from Australia) from the wave 1 main sample that were successfully recontacted at waves 2, 3, 4 and 5. These inflation weights are the basis of the 3 set of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 1–5 longitudinal inflation weights are the waves 1–4 longitudinal inflation weights (computed section 4.2.1) adjusted for attrition between waves 4 and 5. As with the waves 1–4 longitudinal inflation weights, computation of the waves 1–5 longitudinal inflation weights followed the same steps as described in section 1.1.2. Since these weights are meant to make these 4509 respondents representative of the population at the time of wave 1 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 1 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 5 cross-sectional weights for the main

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette smokers					
Cigarette only	242	242	273	228	160
Dual users	89	82	82	64	26
Total	331	324	355	292	186
Exclusive NVP/HTP users	0	0	0	0	0
Quitters					
NVP/HTP users	0	0	0	0	0
Non users	0	0	0	0	0
Total	0	0	0	0	0
Total	331	324	355	292	186

† Variables `kUserGroup` in the dataset

Table 22: Wave 1 respondents successfully recontacted at wave 5 by user group.

sample (see section 4.1.2). Note that respondents were divided into the 4 user groups based on their wave 1 status (variables `kUserGroup`), and that age at wave 1 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Though computation followed the same steps as described in section 1.1.2, too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 5, and it was decided to no longer compute waves 1–5 longitudinal weights for those user groups.

1- Variable `oWTS221v` contains the rescaled waves 1–5 longitudinal weights for the 3498 respondents (873 from the US, 1172 from Canada, 837 from England and 616 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 1 data collection and successfully recontacted at wave 5. These are simply the waves 1–5 longitudinal inflation weights (computed above) of those 3498 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 1 data collection.

2- Variable `oWTS421v` contains the rescaled waves 1–5 longitudinal weights for the 860 respondents (170 from the US, 346 from Canada, 248 from England and 96 from Australia) who were dual users at the time of wave 1 data collection and successfully recontacted at wave 5. These are simply the waves 1–5 longitudinal inflation weights (computed above) of those 860 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 1 data collection.

3- Variable `oWTS121v` contains the rescaled waves 1–5 longitudinal weights for all 4509 respondents (1256 from the US, 1443 from Canada, 1047 from England and 763 from Australia) from the wave 1 main sample that were successfully recontacted at wave 5. As mentioned at the top of section 5.2 and in table 20, too few of the exclusive NVP users and quitters recruited at wave 1 were retained and interviewed at wave 5, and it was decided to no longer compute waves 1–5 longitudinal weights for those user groups. Consequently, variable `oWTS121v` is the same as `oWTS221v`.

5.2.2 Waves 1–5 longitudinal weights for the US reduced sample

The following 3 sets of waves 1–5 longitudinal weights were computed for the 1225 respondents from the wave 1 US reduced sample (see third column of table 22) that were successfully recontacted at waves 2, 3, 4 and 5. Computation followed the same steps as described in section 5.2.1, and the following weights are also designed to make those 1225 respondents representative at the time of wave 1.

As in section 5.2.1, no weights were computed for exclusive NVP users and quitters recruited at wave 1 and retained/interviewed at wave 5, as there were simply too few of them to be representative.

1- Variable `oWTS261v` is the analogue of variable `oWTS221v`, but for the reduced US sample of 850 respondents who were cigarette smokers at the time of wave 1 data collection.

2- Variable `oWTS461v` is the analogue of variable `oWTS421v`, but for the reduced US sample of 164 respondents who were dual users at the time of wave 1 data collection.

3- Variable `oWTS161v` is the analogue of variable `oWTS121v`, but for the reduced US sample of 1225 respondents. In the present case, it is the same as variable `oWTS261v`, since no waves 1–5 longitudinal weights were computed for exclusive NVP users and quitters.

5.3 Waves 2–5 longitudinal sampling weights

As mentioned at the beginning of section 5, 42 sets of longitudinal weights were computed at wave 5 of the 4CV Survey. Twelve of those (see middle quarter of table 20) were computed for respondents interviewed at wave 2, and then successfully retained and interviewed at waves 3, 4 and 5. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 2—5 longitudinal weights are designed to make the sample representative at the time of wave 2. If the aim is to ensure that the sample is representative at the time of wave 5 data collection, then the wave 5 cross-sectional sampling weights computed in section 5.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

As mentioned in table 20, too few Canadian, British and Australian exclusive NVP users recruited at waves 1 or 2 were retained and interviewed at wave 5, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–5 longitudinal weights those individuals.

5.3.1 Waves 2–5 longitudinal weights for the main sample

Waves 2–5 longitudinal inflation weights were computed for the 4509 respondents (1256 from the US, 1443 from Canada, 1047 from England and 763 from Australia) from the wave 2 main sample that were successfully recontacted at waves 3, 4 and 5. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 2–5 longitudinal inflation weights are the waves 2–4 longitudinal inflation weights (computed in section 4.3.1) adjusted for attrition between waves 4 and 5. Consequently, computation of those

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette smokers					
Cigarette only	426	416	364	211	212
Dual users	133	127	120	64	43
Total	559	543	484	275	255
Exclusive NVP/HTP users	26	24	0	0	0
Quitters					
NVP/HTP users	70	69	29	0	21
Non users	22	22	24	0	13
Total	92	91	53	0	34
Total	677	658	537	275	289

† Variables `lUserGroup` and `lQuitGroup` in the dataset

Table 23: Wave 2 respondents successfully recontacted at wave 5 by user group.

longitudinal inflation weights followed the same steps as described in section 2.1.1. Since these weights are meant to make these 4509 respondents representative of the population at the time of wave 2 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 2 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 5 cross-sectional weights for the main sample (see section 5.1.1). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 2 status (variables `lUserGroup` and `lQuitGroup`), and that age at wave 2 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

Though computation followed the same steps as described in section 2.1.1, too few Canadian, British and Australian exclusive NVP users from the wave 2 main sample were retained and interviewed at wave 5, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–5 longitudinal weights those individuals.

1- Variable `oWTS223v` contains the rescaled waves 2–5 longitudinal weights for the 3498 respondents (873 from the US, 1172 from Canada, 837 from England and 616 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 2 data collection and successfully recontacted at waves 3, 4 and 5. These are simply the waves 2–5 longitudinal inflation weights (computed above) of those 3498 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 2 data collection.

2- Variable `oWTS323v` contains the rescaled waves 2–5 longitudinal weights for the 1423 respondents (459 from the US, 440 from Canada, 364 from England and 160 from Australia) who were at least monthly NVP users at the time of wave 2 data collection and successfully recontacted at waves 3, 4 and 5. These are simply the waves 2–5 longitudinal inflation weights (computed above) of those 1423 respondents rescaled to sum to sample size in each country.

Cautionary note: As mentioned at the beginning of section 5.3 and in table 23, too few Canadian, British and Australian exclusive NVP users recruited at waves 1 or 2 were retained and interviewed at wave 5, and likewise for British quitters. Since the representativeness of such small samples is highly questionnable, it was decided to no longer compute waves 2–5 longitudinal weights those individuals (thus no `oWTS323v` weights were computed for those individuals). This is of key importance, as this means that the target population for the `oWTS323v` weights is not the same for all 4 countries. That is, these weights are designed to make American NVP users representative of the population of at least monthly NVP users. For Canada and Australia, these weights are designed to make NVP users representative of the population consisting of dual users and quitters who use NVPs at least monthly. For England, these weights are designed to make NVP users representative of the population consisting of dual users.

3- Variable `oWTS423v` contains the rescaled waves 2–5 longitudinal weights for the 860 respondents (170 from the US, 346 from Canada, 248 from England and 96 from Australia) who were dual users at the time of wave 2 data collection and successfully recontacted at waves 3, 4 and 5. These are simply the waves 2–5 longitudinal inflation weights (computed above) of those 860 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 2 data collection.

4- Variable `oWTS523v` contains the rescaled waves 2–5 longitudinal weights for the 764 respondents (247 from the US, 233 from Canada, 161 from England and 123 from Australia) who had quit smoking cigarettes at the time of wave 2 data collection and successfully recontacted at waves 3, 4 and 5. These are simply the waves 2–5 longitudinal inflation weights (computed above) of those 764 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country (with the exception of England obviously[21]) at the time of wave 2 data collection. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

5- Variable `oWTS623v` contains the rescaled waves 2–5 longitudinal weights for the 4061 respondents (1162 from the US, 1266 from Canada, 953 from England and 680 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection and successfully recontacted at waves 3, 4 and 5. These are simply the waves 2–5 longitudinal inflation weights (computed above) of those 4061 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 2 data collection.

The cautionary note about exclusive NVP users and English quitters stated with variable `oWTS323v` also applies here.

6- Variable `oWTS123v` contains the rescaled waves 2–5 longitudinal weights for all 4509 respondents (1256 from the US, 1443 from Canada, 1047 from England and 763 from Australia) from the wave 2 main sample that were successfully recontacted at waves 3, 4 and 5. These are simply the waves 2–5 longitudinal inflation weights (computed above) of those 4509 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1443 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users

[21] Recall that too few British quitter recruited at waves 1 or 2 were retained and interviewed at wave 5. Consequently, it was decided to no longer compute waves 2–5 longitudinal for those individuals.

and quitters (within the last 5 years) at the time of wave 2 data collection; likewise for the US and Australia (within the last 2 years). For England, these weights are designed to make these 1047 British tobacco users representative of the British population of cigarette smokers and dual users.

The cautionary note about exclusive NVP users and English quitters stated with variable oWTS323v also applies here.

As mentioned in the description of variable oWTS101v, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the nTS123v weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

5.3.2 Waves 2–5 longitudinal weights for the US reduced sample

The following 6 sets of waves 2–5 longitudinal weights were computed for the 1225 respondents from the wave 2 US reduced sample (see third column of table 23) that were successfully recontacted at waves 3, 4 and 5. Computation followed the same steps as described in section 5.3.1, and the following weights are also designed to make those 1225 respondents representative at the time of wave 2.

1- Variable oWTS263v is the analogue of variable oWTS223v, but for the reduced US sample of 850 respondents who were cigarette smokers at the time of wave 2 data collection.

2- Variable oWTS363v is the analogue of variable oWTS323v, but for the reduced US sample of 448 respondents who were at least monthly NVP users at the time of wave 2 data collection.

3- Variable oWTS463v is the analogue of variable oWTS423v, but for the reduced US sample of 164 respondents who were dual users at the time of wave 2 data collection.

4- Variable oWTS563v is the analogue of variable oWTS523v, but for the reduced US sample of 242 respondents who were quitters at the time of wave 2 data collection.

5- Variable oWTS663v is the analogue of variable oWTS623v, but for the reduced US sample of 1134 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 2 data collection.

6- Variable oWTS163v is the analogue of variable oWTS123v, but for the reduced US sample of 1225 respondents.

5.4 Waves 3–5 longitudinal sampling weights

As mentioned at the beginning of section 5, 42 sets of longitudinal weights were computed at wave 5 of the 4CV Survey. Twelve of those (see middle quarter of table 20) were computed for respondents interviewed at wave 3, and then successfully retained and interviewed at waves 4 and 5. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 3—5 longitudinal weights are designed to make the sample representative at the time of wave 3. If the aim is to ensure that the sample is representative at the time of wave 5 data collection, then the wave 5 cross-sectional sampling weights computed in section 5.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for

an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette smokers					
Cigarette only	547	537	547	209	307
Dual users	136	131	198	50	70
Total	683	668	745	259	377
Exclusive NVP/HTP users	110	108	13	14	20
Quitters					
NVP/HTP users	76	74	40	24	31
Non users	76	76	67	24	31
Total	152	150	107	48	62
Total	945	926	865	321	459

† Variables `mUserGroup` and `mQuitGroup` in the dataset

Table 24: Wave 3 respondents successfully recontacted at wave 5 by user group.

5.4.1 Waves 3–5 longitudinal weights for the main sample

Waves 3–5 longitudinal inflation weights were computed for the 2590 respondents (945 from the US, 865 from Canada, 321 from England and 459 from Australia) from the wave 3 main sample that were successfully recontacted at wave 5. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 3–5 longitudinal inflation weights are the waves 3–4 longitudinal inflation weights (computed in section 4.4.1) adjusted for attrition between waves 4 and 5. Consequently, computation of those longitudinal inflation weights followed the same steps as described in section 3.1.2. Since these weights are meant to make these 2590 respondents representative of the population at the time of wave 3 data collection, the national benchmark surveys (i.e., NHIS, CTADS, NDSHS, OLS and STS) used to compute the wave 3 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 5 cross-sectional weights for the main sample (see section 5.1.1). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 3 status (variables `mUserGroup` and `mQuitGroup`), and that age at wave 3 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

1- Variable `oWTS225v` contains the rescaled waves 3–5 longitudinal weights for the 2064 respondents (683 from the US, 745 from Canada, 259 from England and 377 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 3 data collection and

successfully recontacted at wave 5. These are simply the waves 3–5 longitudinal inflation weights (computed above) of those 2064 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 3 data collection.

2- Variable `oWTS325v` contains the rescaled waves 3–5 longitudinal weights for the 782 respondents (322 from the US, 251 from Canada, 88 from England and 121 from Australia) who were at least monthly NVP users at the time of wave 3 data collection and successfully recontacted at wave 5. These are simply the waves 3–5 longitudinal inflation weights (computed above) of those 782 respondents rescaled to sum to sample size in each country. These weights are designed to make these NVP users representative of the population of at least monthly NVP users of their respective country at the time of wave 3 data collection.

3- Variable `oWTS425v` contains the rescaled waves 3–5 longitudinal weights for the 454 respondents (136 from the US, 198 from Canada, 50 from England and 70 from Australia) who were dual users at the time of wave 3 data collection and successfully recontacted at wave 5. These are simply the waves 3–5 longitudinal inflation weights (computed above) of those 454 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 3 data collection.

4- Variable `oWTS525v` contains the rescaled waves 3–5 longitudinal weights for the 369 respondents (152 from the US, 107 from Canada, 48 from England and 62 from Australia) who had quit smoking cigarettes (within the last 5 years) at the time of wave 3 data collection and successfully recontacted at wave 5. These are simply the waves 3–5 longitudinal inflation weights (computed above) of those 369 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country at the time of wave 3 data collection. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

5- Variable `oWTS625v` contains the rescaled waves 3–5 longitudinal weights for the 2392 respondents (869 from the US, 798 from Canada, 297 from England and 428 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 3 data collection and successfully recontacted at wave 5. These are simply the waves 3–5 longitudinal inflation weights (computed above) of those 2392 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 3 data collection.

6- Variable `oWTS125v` contains the rescaled waves 3–5 longitudinal weights for all 2590 respondents (945 from the US, 865 from Canada, 321 from England and 459 from Australia) from the wave 3 main sample that were successfully recontacted at wave 5. These are simply the waves 3–5 longitudinal inflation weights (computed above) of those 2590 respondents rescaled to sum to sample size in each country. These weights are designed to make these 865 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 3 data collection; likewise for the US, England and Australia.

   As mentioned in the description of variable `oWTS101v`, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the `oWTS125v` weights. This is probably fine when the goal is to carry

out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

5.4.2 Waves 3–5 longitudinal weights for the US reduced sample

The following 6 sets of waves 3–5 longitudinal weights were computed for the 926 respondents from the wave 3 US reduced sample (see third column of table 24) that were successfully recontacted at waves 4 and 5. Computation followed the same steps as described in section 5.4.1, and the following weights are also designed to make those 926 respondents representative at the time of wave 3.

1- Variable `oWTS265v` is the analogue of variable `oWTS225v`, but for the reduced US sample of 668 respondents who were cigarette smokers at the time of wave 3 data collection.

2- Variable `oWTS365v` is the analogue of variable `oWTS325v`, but for the reduced US sample of 313 respondents who were at least monthly NVP users at the time of wave 3 data collection.

3- Variable `oWTS465v` is the analogue of variable `oWTS425v`, but for the reduced US sample of 131 respondents who were dual users at the time of wave 3 data collection.

4- Variable `oWTS565v` is the analogue of variable `oWTS525v`, but for the reduced US sample of 150 respondents who were quitters at the time of wave 3 data collection.

5- Variable `oWTS665v` is the analogue of variable `oWTS625v`, but for the reduced US sample of 850 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 3 data collection.

6- Variable `oWTS165v` is the analogue of variable `oWTS125v`, but for the reduced US sample of 926 respondents.

5.5 Waves 4–5 longitudinal sampling weights

As mentioned at the beginning of section 5, 42 sets of longitudinal weights were computed at wave 5 of the 4CV Survey. Twelve of those (see bottom quarter of table 20) were computed for respondents interviewed at wave 4, and then successfully retained and interviewed at wave 5. This section details the computation of those weights, as well as how and when they should be used. It should be noted that all waves 4—5 longitudinal weights are designed to make the sample representative at the time of wave 4. If the aim is to ensure that the sample is representative at the time of wave 5 data collection, then the wave 5 cross-sectional sampling weights computed in section 5.1 should be used instead. Please note that the comment, at the beginning of section 2.2.2, about the possibility for an individual to have a longitudinal weight though he/she does not have a cross-sectional weight also applies here.

5.5.1 Waves 4–5 longitudinal weights for the main sample

Waves 4–5 longitudinal inflation weights were computed for the 3858 respondents (1487 from the US, 1227 from Canada, 437 from England and 707 from Australia) from the wave 4 main sample that were successfully recontacted at wave 5. These inflation weights are the basis of the 6 sets of weights described below. However, they are not part of the standard 4CV dataset, as the vast majority of longitudinal analyses will focus on analytical inference (as opposed to descriptive inference) and that the rescaled

User group†	US		Canada	England	Australia
	Main	Reduced			
Cigarette smokers					
Cigarette only	816	804	621	228	416
Dual users	204	201	354	104	135
Total	1020	1005	975	332	551
Exclusive NVP/HTP users	145	141	35	28	24
Quitters					
NVP/HTP users	143	143	85	32	56
Non users	179	179	132	45	76
Total	322	322	217	77	132
Total	1487	1468	1227	437	707

† Variables `nUserGroup` and `nQuitGroup` in the dataset

Table 25: Wave 4 respondents successfully recontacted at wave 5 by user group.

weights are preferable in such cases; see section 6.2 for more information on inflation versus rescaled weights. Nonetheless, if needed, these inflation weights are available by contacting the first author of this report.

These waves 4–5 longitudinal inflation weights are the wave 4 cross-sectional inflation weights (i.e., variable `nWTS100v`, computed in section 4.1.2) adjusted for attrition between waves 4 and 5. Consequently, computation of those longitudinal inflation weights followed the same steps as described in section 4.1.2. Since these weights are meant to make these 3858 respondents representative of the population at the time of wave 4 data collection, the national benchmark surveys (i.e., NHIS, CCHS, NDSHS, OLS and STS) used to compute the wave 4 cross-sectional weights were used here as well; as opposed to the updated figures used to compute the wave 5 cross-sectional weights for the main sample (see section 5.1.1). Note that respondents were divided into the 4 user groups and 4 quitter subgroups based on their wave 4 status (variables `nUserGroup` and `nQuitGroup`), and that age at wave 4 was used when dividing respondents into the different age groups. Likewise for ethnicity, education, language and geographic region.

1- Variable `oWTS227v` contains the rescaled waves 4–5 longitudinal weights for the 2878 respondents (1020 from the US, 975 from Canada, 332 from England and 551 from Australia) who were cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) at the time of wave 4 data collection and successfully recontacted at wave 5. These are simply the waves 4–5 longitudinal inflation weights (computed above) of those 2878 respondents rescaled to sum to sample size in each country. These weights are designed to make these cigarette smokers representative of the population of cigarette smokers (and have smoked at least 100 cigarettes in their lifetime) of their respective country at the time of wave 4 data collection.

2- Variable `oWTS327v` contains the rescaled waves 4–5 longitudinal weights for the 1345 respondents (492 from the US, 474 from Canada, 164 from England and 215 from Australia) who were at least monthly NVP users at the time of wave 4 data collection and successfully recontacted at wave 5. These are simply the waves 4–5 longitudinal inflation weights (computed above) of those 1345 respondents

rescaled to sum to sample size in each country. These weights are designed to make these NVP users representative of the population of at least monthly NVP users of their respective country at the time of wave 4 data collection.

3- Variable `oWTS427v` contains the rescaled waves 4–5 longitudinal weights for the 797 respondents (204 from the US, 354 from Canada, 104 from England and 135 from Australia) who were dual users at the time of wave 4 data collection and successfully recontacted at wave 5. These are simply the waves 4–5 longitudinal inflation weights (computed above) of those 797 respondents rescaled to sum to sample size in each country. These weights are designed to make these dual users representative of the population of dual users of their respective country at the time of wave 4 data collection.

4- Variable `oWTS527v` contains the rescaled waves 4–5 longitudinal weights for the 748 respondents (322 from the US, 217 from Canada, 77 from England and 132 from Australia) who had quit smoking cigarettes (within the last 5 years) at the time of wave 4 data collection and successfully recontacted at wave 5. These are simply the waves 4–5 longitudinal inflation weights (computed above) of those 748 respondents rescaled to sum to sample size in each country. These weights are designed to make these quitters representative of the population of quitters of their respective country at the time of wave 4 data collection. The cautionary notes of variable `kWTS501v` and of section 6.3 still apply here.

5- Variable `oWTS627v` contains the rescaled waves 4–5 longitudinal weights for the 3426 respondents (1308 from the US, 1095 from Canada, 392 from England and 631 from Australia) who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 4 data collection and successfully recontacted at wave 5. These are simply the waves 4–5 longitudinal inflation weights (computed above) of those 3426 respondents rescaled to sum to sample size in each country. These weights are designed to make these tobacco users representative of the population of tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) of their respective country at the time of wave 4 data collection.

6- Variable `oWTS127v` contains the rescaled waves 4–5 longitudinal weights for all 3858 respondents (1487 from the US, 1227 from Canada, 437 from England and 707 from Australia) from the wave 4 main sample that were successfully recontacted at wave 5. These are simply the waves 4–5 longitudinal inflation weights (computed above) of those 3858 respondents rescaled to sum to sample size in each country. These weights are designed to make these 1227 Canadian tobacco users and quitters (within the last 5 years) representative of the Canadian population of tobacco users and quitters (within the last 5 years) at the time of wave 4 data collection; likewise for the US, England and Australia.

   As mentioned in the description of variable `oWTS101v`, it should be noted that tobacco users and quitters are ultimately two distinct populations. Hence, great care must be taken when deciding to analyse them together using the `oWTS125v` weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users and quitters. However, carrying out analytical inference (which is likely to be of interest in a longitudinal analysis) from that same joint population is probably much more questionable.

5.5.2 Waves 4–5 longitudinal weights for the US reduced sample

The following 6 sets of waves 4–5 longitudinal weights were computed for the 1468 respondents from the wave 4 US reduced sample (see third column of table 24) that were successfully recontacted at wave 5. Computation followed the same steps as described in section 5.5.1, and the following weights are also designed to make those 1468 respondents representative at the time of wave 4.

1- Variable `oWTS267v` is the analogue of variable `oWTS227v`, but for the reduced US sample of 1005 respondents who were cigarette smokers at the time of wave 4 data collection.

2- Variable `oWTS367v` is the analogue of variable `oWTS327v`, but for the reduced US sample of 485 respondents who were at least monthly NVP users at the time of wave 4 data collection.

3- Variable `oWTS467v` is the analogue of variable `oWTS427v`, but for the reduced US sample of 201 respondents who were dual users at the time of wave 4 data collection.

4- Variable `oWTS567v` is the analogue of variable `oWTS527v`, but for the reduced US sample of 322 respondents who were quitters at the time of wave 4 data collection.

5- Variable `oWTS667v` is the analogue of variable `oWTS627v`, but for the reduced US sample of 1289 respondents who were tobacco users (i.e., cigarette smokers and/or at least monthly NVP users) at the time of wave 4 data collection.

6- Variable `oWTS167v` is the analogue of variable `oWTS127v`, but for the reduced US sample of 1468 respondents.

6 Remarks and cautionary notes

This section provides remarks and cautionary notes that must be followed when analyzing 4CV data. These echo some of the limitations of the 4CV Survey. In addition to this section, users should also look at the cautionary notes in sections 1.1.1 and 3.1.1.

6.1 Estimating prevalences

Although it offers rich possibilities for analyses, the 4CV Survey is not suited for estimating prevalences of cigarette smoking, use of NVP, dual use and quitting. Specifically, it is not suited for estimating prevalences/proportions of the following five user groups: *i*) cigarette only users, *ii*) dual users, *iii*) exclusive NVP users and *iv*) quitters using NVP and *v*) quitters not using NVP; see description in sections 1.1.1 and 3.1.1. This is largely due to its sampling design, where 18—24 years old and users of NVP are oversampled, and to how the sampling weights were computed to account for this sampling design and other biases.

The sampling design of the 4CV was chosen to ensure sufficiently large sample sizes for those two types of respondents, while keeping the overall sample size and cost reasonable. The downside of such a sampling design, in particular the oversampling of NVP users, is that it requires the use of external benchmark/calibration data to compute sampling weights. Note that the same benchmark/calibration figures were also used to reduce various biases in the sample. As described earlier (see, for example, the description of variable `kWTS100v` in section 1.1.2) various national health surveys were used to obtain the required benchmark/calibration figures; see appendix A.2. A raking algorithm (see appendix A.3) was then used to compute the weights of the 4CV Survey based on these figures.

The consequence/drawback of using external benchmark/calibration figures to compute the weights is that the various prevalence estimates obtained from the 4CV data are actually those of the various national health surveys. For example, the estimated prevalence of dual use obtained from US respondents to the 4CV is actually that of the National Health Interview Survey (NHIS) (the national health survey used to benchmark the US sample); likewise for the prevalences/proportions of the other 4 user groups.

Because the weights were calibrated on user group × gender, user group × age group, and user group × education, this even holds when looking at prevalences within gender, age group and education.

Lastly, because of rapid change in patterns of use (especially with respect to the use of NVP), the calibration figures may not correspond well to the pattern which existed at the time the 4CV data was collected.

6.2 Inflation versus rescaled weights

The main reason for rescaling the weights is to facilitate joint analyses involving data from multiple ITC countries. From the data used to calibrate the weights (see appendix A.2), there were about 38.1 million cigarette smokers in the United States at the time of wave 1 data collection; whereas there were only 3.6 millions such individuals in Canada, 7.6 millions in England and 2.6 millions in Australia. Hence, any joint analysis using data from all four countries will be dominated by the US if the inflation weights (e.g., variables `kWTS100v` or `kWTS102v`). Though the number of smokers changed a little at waves 2, 3, 4 and 5, this also holds true at those waves. Hence, analyses using inflation weights (e.g., variables `lWTS100v`, `lWTS102v`, `mWTS100v`, `mWTS102v`, `nWTS100v`, `nWTS102v`, `oWTS100v` or `oWTS102v`) will also be dominated by the US.

On the other hand, the various rescaled weights sum to their corresponding sample size. Hence, if the rescaled weights are used, England and Canada will have a slightly greater impact on the results (since the Canadian and British sample sizes are larger than the Australian and US samples; see tables 3, 6, 10, 15 and 21), but no country will dominate the analysis. In summary, rescaling the weights to sum to the sample size is a simple and efficient way to make countries with different population sizes comparable. This also holds true when comparing 4CV data to other ITC countries; for examples, ITC Netherlands, ITC Japan, ITC Korea and ITC 6E.

Last but not least, it should be mentioned that rescaling the weights will not affect the results when estimating population means and proportions/percentages, as well as when fitting various statistical models (e.g., logistic and linear regressions). However, the rescaled weights should not be used to estimate population totals (e.g., the total number of daily smokers or NVP users).

6.3 Cautionary notes about quitters

As mentioned earlier (see, for example, variable `kWTS501v`), many quitters in the 4CV sample were initially recruited as smokers. Since quitting is one of the reasons for dropping out of the survey, the quitting experienced by our cohort respondents could well be affected by being in the sample and because of the sampling design itself, quitters in the 4CV sample should not be considered to be representative of quitters in the population. For example, comparisons between the quitters in the ITC sample and quitters in the cross-sectional Smoking Toolkit Study showed an important discrepancy in the distribution of length of time quit.

As described in sections 1.1.2, 2.1.1, 3.1.2, 4.1.2 and 5.1.1 (see variables `kWTS100v`, `lWTS100v`, `mWTS100v`, `nWTS100v` and `oWTS100v`), the sampling weights of quitters were calibrated on gender, age group, geographic region, ethnicity (US only), education (except for Canada), language (Canada only) and use of NVP × length of quit (≤ 1 year vs. > 1 year; though length of quit was omitted on occasions due to small sample sizes). Quitters in the 4CV sample should thus be representative of the population in terms of those variables, but not necessarily in terms of other attributes. Again, they are an imperfect sample when it comes being representative of the whole population of quitters.

6.4 Cautionary notes about NVP users

Comparisons on measures related to NVP between the 4CV Survey and other ITC countries must be viewed with caution, as weight construction for the 4CV Survey was done in a different fashion than that of other ITC countries.

As described in sections 1.1.2, 2.1.1, 3.1.2, 4.1.2 and 5.1.1 (see variables kWTS100v, lWTS100v, mWTS100v, nWTS100v and oWTS100v), weights were computed by first dividing respondents into four broad user groups: cigarette only users, dual users, exclusive NVP users and quitters. The sampling weights of dual users and of quitters were then calibrated on gender, age group, geographic region, ethnicity (US only), education (except for Canada waves 1–3) and language (Canada only waves 1, 2 and 4), whereas those who exclusively use NVPs were calibrated on gender and age group. The sampling weights of quitters were also calibrated on use of NVPs and length of quit (≤ 1 year vs. > 1 year; though length of quit was omitted on occasions due to small sample sizes). In some of the other ITC countries (e.g., Japan and Korea) the weights are computed in a similar fashion. In most of the other ITC countries however, separate estimates for the number of individuals who are dual users and individuals who only smoke cigarettes were not available at the time of weight calculation. Hence, the weights were calibrated using smoking prevalences (often by age/gender groups and/or geographic regions), and thus have no special adjustment for NVP usage.

6.5 Covariates to include in statistical modelling

As with other surveys, it is good practice to include the survey design variables and the variables used in the weight construction, when fitting statistical models (e.g., linear or logistic regression models) using 4CV data. Hence, we highly recommend that any statistical model includes the following covariates: gender, age (either as a continuous or as a categorical covariate) and user group (see description of variables kWTS100v, lWTS100v, mWTS100v, nWTS100v or oWTS100v). The geographic region should also be used as the stratification variable in the statistical software. Table 26 contains the variable names of those covariates. Though somewhat less essential, users should also strongly consider adding the following covariates:

	Variable names				
Covariates	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
gender			gender		
ethnicity			ethnic		
age (continuous)	kAGE	lAGE	mAGE	nAGE	oAGE
age (categorical)	kageGroup	lageGroup	mageGroup	nageGroup	oageGroup
user group	kUserGroup	lUserGroup	mUserGroup	nUserGroup	oUserGroup
education	kDE312v	lDE312v	mDE312v	nDE312v	oDE312v
language	kcaFrench	lcaFrench	mcaFrench	ncaFrench	ocaFrench
quit group	kQuitGroup	lQuitGroup	mQuitGroup	nQuitGroup	oQuitGroup
vaping frequency	kEC309v	lEC309v	mEC309v	nEC309v	oEC309v
strata/geographic region	kStrata	lStrata	mStrata	nStrata	oStrata

Table 26: Covariates to include in statistical models

- ethnicity, when fitting models using US data
- education, when fitting models using US, English, Australian and/or wave 4 Canadian data

- language, when fitting models using Canadian data
- use of NVP and/or length of quit, when fitting data using quitters; see description of the kWTS100v, lWTS100v, mWTS100v, nWTS100v and oWTS100v weight variables
- frequency of use of NVP, when fitting models with NVP users.

Funding & acknowledgements

This study was supported by grants from the U.S. National Cancer Institute (P01CA200512), the Canadian Institutes of Health Research (FDN-148477), and by the National Health and Medical Research Council of Australia (APP 1106451).

The authors would like to acknowledge and thank all those that contributed to the International Tobacco Control Four Country Smoking and Vaping (4CV) Survey: all study investigators and collaborators, and the project staff at their respective institutions.

Short description for published papers

Below is a short description/summary of the various weight constructed in this report. This description aims to help authors provide a summary of the weight calculation process of the 4CV.

For the US, Canada and Australia, base cross-sectional sampling weights were computed using the method of Chen, Li & Wu (2020). A raking algorithm was then used to calibrate these weights to benchmark population figures on smoking and vaping obtained from the national health surveys (U.S. National Health Interview Survey (NHIS), Canadian Community Health Survey (CCHS) and the Australian National Drug Strategy Household Survey (NDSHS)). This calibration was done based on smoking status (including quit status and length of quit) and vaping status, gender, age, geographic regions, education, ethnicity (US only), and language (Canada only). For England, due to limitation in data access, it was not possible to use the method of Chen, Li & Wu, and every respondent was assigned a base cross-sectional sampling weights of 1. As with the other 3 countries, a raking algorithm was then used to calibrate these weights to benchmark population figures on smoking and vaping obtained from the Opinions and Lifestyle Survey (OLS). This calibration was done based on smoking status and vaping status, gender, age, geographic regions and education.

Many sets of longitudinal sampling weights were also computed for respondents retained over multiple waves. These weights adjusted for attrition between two consecutive waves, and used the same raking algorithm and national health surveys as described above for the cross-sectional weights.

References

Chen, Y., Li, P. & Wu, C. (2020), 'Doubly robust inference with nonprobability survey samples', *Journal of the American Statistical Association* **115**, 2011–2021.

A Appendix

A.1 Pseudo code

The following pseudo code details how variables `UserGroup` (i.e., variable `kUserGroup` at wave 1, `lUserGroup` at wave 2, `mUserGroup` at wave 3, `nUserGroup` at wave 4, etc.) and `QuitGroup` (i.e., variable `kQuitGroup` at wave 1, `lQuitGroup` at wave 2, `mQuitGroup` at wave 3, `nQuitGroup` at wave 4, `oQuitGroup` at wave 5, etc.) were created for US, Canada and England. The code for Australia is identical, with the exception that cut-off for quitters is 2 years instead of 5 at waves 1–2 (see section 3.1.1). Note that the 4CV survey questionnaires are available here, and contain the actual questions and response options associated with the variables `FR309v`, `EC309v` and `QA439`.

```
define
A - current cigarette user- smokes cigarette at least monthly or occasionally
                                                    (ie, FR309v in [10,20,31,32])
B - current e-cigarette user- uses e-cigarette at least once a month  (ie, EC309v in [10,20,31])
C - current NVP user- uses NVP at least once a month  (ie, HN309v in [10,20,31])
D - current ONP user - uses ONP at least once a month (ie, NC006 in (1,2,3) or BI038 in (1,2,3))
E - current cigarettes quitter - was cigarettes user, but does not smoke
                                                      at current wave (ie, FR309v in [40,80])
F - never smokes cigarettes or smokes less than 100 cigarettes (ie, FR309v=90)

Wave 1-3
if { A and not B and not C } then UserGroup=1; /* cig only */;
else if { A and (B or C) } then UserGroup=2; /* dual user */;
else if { F and kQA439=10/lQA439=11/mQA439=11 and (B or C) see * below) }
                                                  then UserGroup=3; /* exclusive NVP */;
else if { E } then UserGroup=4; /* quitter */;

Wave 4-5

if { A and not B and not C and not D } then UserGroup=1; /* cig only */;
else if { A and (B or C or D) } then UserGroup=2; /* dual user */;
else if { F and kQA439=10/lQA439=11/mQA439=11 and (B or C or D) see * below) }
                                                    then UserGroup=3; /* exclusive NVP */;
else if { E } then UserGroup=4; /* quitter */;


if {quitters (ie, UserGroup=4)} then do;
  if {quit within the last 12 months (ie, QA439<=6)}
      and { B or C or D }
      then QuitGroup=1;

  else if {quit 1-5 years ago (ie, 6<kQA439<10/6<lQA439/mQA439/nQA439/oQA439<11; see * below)}
       and { B or C or D }
       then QuitGroup=2;

  else if {quit within the last 12 months (ie, QA439<=6)}
       and { not B and not C and not D }
       then QuitGroup=3;

  else if {quit 1-5 years ago (ie, 6<kQA439<10/6<lQA439<11; see * below)}
       and { not B and not C and not D }
       then QuitGroup=4;
end;
```

* The response options for variable `QA439` were changed at wave 2, and this is why the code for wave 1 (i.e., variable `kQA439`) is a little different than that for wave 2 (i.e., variable `lQA439`) and for wave 3 (i.e., variable `mQA439`).

A.2 Benchmark/calibration figures

A.2.1 Canada

The benchmark/calibration figures used to compute the weights for Canadian respondents are given below. The 2015 Canadian Tobacco Alcohol and Drugs Survey (CTADS) was used for the wave 1 weights (as well as for all longitudinal weights computed for the wave 1 sample successfully recontacted at follow-up waves; i.e., waves 1–2, waves 1–3, waves 1–4, etc.), whereas the 2017 CTADS was used for the wave 2 cross-sectional weights (as well as for all longitudinal weights computed for the wave 2 sample successfully recontacted at follow-up waves; i.e., waves 2–3, waves 2–4, waves 2–5, etc.). The 2019 Canadian Tobacco and Nicotine Survey (CTNS) was used for the wave 3 cross-sectional weights (as well as the waves 3–4 and 3–5 longitudinal weights). Please note that CTNS 2019 data available at the time of weight calculation did not allow for obtaining benchmark/calibration figures for individuals aged 18–19 years old. To compensate for this, CTADS 2017 data for this age group was used to obtain an estimate. Lastly, the 2021 Canadian Community Health Survey (CCHS) was used for the wave 4 cross-sectional weights (as well as the waves 4–5 longitudinal weights), and the 2022 (CCHS) was used for the wave 5 cross-sectional weights

As mentioned at the end of section 3.1.1, the 2019 CTNS provides little information to properly estimate the number of HTP users at wave 3. Hence, some (in particular exclusive NVP) of the benchmark figures (in the second to last column of the tables below) are likely to somewhat underestimate the true values. Like the 2019 CTSN, the 2021 and 2022 CCHS provides little information to properly estimate the number of HTP. This is a little better for snus and of tobacco-free oral nicotine products products (ONPs). Nevertheless, the waves 4 and 5 benchmark figures suffer from simular issues as the wave 3 ones.

		# Individuals				
User group	Sex	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	male	1,893,352	1,704,836	1,479,110	1,558,015	1,779,815
	female	1,211,671	1,516,168	1,444,812	1,332,813	1,361,980
dual users	male	256,490	151,400	206,090	257,745	368,204
	female	232,294	127,694	173,258	143,679	242,039
exclusive NVP	male	97,746	80,003	317,360	348,579	454,643
	female	51,838	41,266	165,860	211,819	374,921
quitters	male	878,689	885,794	719,016	1,028,606	755,101
	female	738,671	430,944	456,049	630,988	506,755

		# Individuals				
User group	Age	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	[18, 25)	343,651	239,735	120,656	91,951	62,518
	[25, 45)	1,248,208	1,257,751	1,077,659	992,109	1,075,853
	[45, 100)	1,513,164	1,723,518	1,725,607	1,806,769	2,003,424
dual users	[18, 25)	122,977	42,106	81,611	78,374	99,069
	[25, 45)	142,777	107,329	148,134	211,910	316,394
	[45, 100)	223,030	129,659	149,603	101,140	194,481
exclusive NVP	[18, 25)	78,413	85,795	246,108	290,127	381,673
	[25, 100)	71,171	35,474	237,112	270,272	447,890
quitters	[18, 25)	175,176	106,931	92,335	131,254	81,368
	[25, 45)	712,492	497,893	579,841	836,070	659,398
	[45, 100)	729,692	711,914	502,889	692,270	521,090

		# Individuals				
User group	Region	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	Maritimes	282,238	240,236	241,997	212,966	270,322
	Quebec	737,184	764,681	718,106	717,824	808,747
	Ontario	1,076,236	1,099,524	1,147,027	1,132,340	1,177,633
	Prairies	683,749	701,034	484,241	512,891	560,148
	BC	325,616	415,529	332,551	314,808	324,945
dual users	Maritimes	36,239	13,381	35,013	31,826	45,935
	Quebec	173,320	83,458	95,788	122,632	162395
	Ontario	134,247	62,699	116,287	127,021	195,351
	Prairies	97,458	38,652	74,920	70,519	130,532
	BC	47,520	80,904	57,340	49,427	76,030
exclusive NVP	Quebec			69,615		
	Ontario	149,584	121,269	204,615	560,398	829,564
	the rest			208,990		
quitters	Maritimes	149,237	81,011	87,975	116,597	92,927
	Quebec	369,844	337,135	268,438	374,466	314,031
	Ontario	629,710	493,441	405,563	639,459	466,999
	Prairies	271,668	228,597	284,448	284,230	219,197
	BC	196,901	176,554	128,641	244,842	168,702

		# Individuals			
User group	Language	Wave 1	Wave 2	Wave 4	Wave 5
cigarette only	English	2,478,289	2,593,272	2,224,360	2,391,502
	French	626,734	627,732	666,469	750,292
dual users	English	352,798	193,298	286,897	453,537
	French	135,986	85,796	114,527	156,706
exclusive NVP	Both/All	149,584	121,269	560,399	829,564
quitters	English	1,298,620	1,020,336	1,328,584	977,247
	French	318,740	296,402	331,010	284,609

User group	Education	# Individuals Wave 4	Wave 5
cigarette only	low	1,278,801	1527161
	moderate	1,152,965	1148274
	high	459,063	466359
dual users	low	187,674	273545
	moderate	161,000	236127
	high	52,751	100571
exclusive NVP	low	297,058	357430
	moderate	141,763	289700
	high	121,577	182434
quitters	low	632,340	501551
	moderate	662,407	494837
	high	364,846	265468

Time since quit	Use of NVP	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
< 1 year	yes	113,486	53,556	206,011	57,395	73,881
1 to 2 years					139,129	45,699
2 to 5 years					97,234	174,989
< 1 year	no	1,073,328	1,007,686	738,876	373,151	255,496
1 to 2 years					233,985	170,282
2 to 5 years					524,449	557,230

A.2.2 US

The benchmark/calibration figures used to compute the weights for US respondents are given below. The 2016 National Health Interview Survey (NHIS) was used for the wave 1 weights (as well as for all longitudinal weights computed for the wave 1 sample successfully recontacted at follow-up waves; i.e., waves 1–2, waves 1–3, waves 1–4, etc.), whereas the 2017 NHIS was used for the wave 2 cross-sectional weights (as well as for all longitudinal weights computed for the wave 2 sample successfully recontacted at follow-up waves; i.e., waves 2–3, waves 2–4, waves 2–5, etc.). Likewise the 2018 NHIS was used for the wave 3 cross-sectional weights (as well as the waves 3–4 and 3–5 longitudinal weights), and the 2022 NHIS was used for the wave 4 cross-sectional weights (as well as the waves 4–5 longitudinal weights). Lastly, the 2023 NHIS (in combination with the 2022 NHIS)[22] was used for the wave 4 cross-sectional weights

As mentioned at the end of section 3.1.1, the 2018 NHIS provides little information to properly estimate the number of HTP users at wave 3. Hence, some (in particular exclusive NVP) of the benchmark figures (in the second last column of the tables below) are likely to somewhat underestimate the true values. This also applies to the 2022 & 2023 NHIS, with the added difficulty that little information is provided about snus and of tobacco-free oral nicotine products products (ONPs) users.

[22]The 2023 NHIS doesn't provide information about length of quit, and the 2022 & 2023 NHIS data were thus combined to obtained the benchmark figures for quitters and exclusive NVP.

		# Individuals				
User group	Sex	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	male	18,476,203	16,766,259	16,643,287	12,796,662	12,118,865
	female	15,424,166	13,954,662	13,984,557	10,900,690	9,811,236
dual users	male	2,294,225	1,927,092	2,044,056	2,685,691	2,989,681
	female	1,914,346	1,482,298	1,271,755	1,478,862	1,807,235
exclusive NVP	male	867,848	881,369	1,459,668	5,992,401	5,863,645
	female	419,182	726,079	964,507	3,241,381	3,097,145
quitters	male	8,189,424	7,612,253	7,156,956	7,767,898	7,444,743
	female	6,401,836	6,203,078	4,700,430	5,762,099	6,101,032

		# Individuals				
User group	Age	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	[18, 25)	3,255,760	2,561,349	1,770,480	677,059	496,908
	[25, 40)	9,916,638	9,214,524	9,235,001	5,159,248	4,748,959
	[40, 55)	10,224,756	8,811,045	9,180,526	7,856,908	6,775,477
	[55, 100)	10,503,215	10,134,003	10,441,837	10,004,137	9,908,757
dual users	[18, 25)	631,973	490,499	501,792	690,229	548,412
	[25, 40)	1,357,264	1,311,808	1,370,997	1,753,362	2,161,661
	[40, 55)	1,314,813	876,076	893,324	1,092,662	1,356,142
	[55, 100)	904,521	731,007	549,698	628,300	730,702
exclusive NVP	[18, 25)	482,790	802,063	1,177,137	7,997,599	2,633,028
	[25, 40)			612,046		3,301,426
	[40, 55)	804,240	805,385	342,634	1,236,183	1,838,442
	[55, 100)			292,358		1,187,895
quitters	[18, 25)	1,333,922	1,380,468	1,362,716	1,338,880	1,553,668
	[25, 40)	5,945,596	5,597,147	4,594,446	5,651,182	5,625,597
	[40, 55)	3,795,597	3,201,832	2,753,171	2,920,605	2,847,384
	[55, 100)	3,516,145	3,635,884	3,147,053	3,619,330	3,519,126

		# Individuals				
User group	Region†	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	Northeast	5,408,604	4,514,875	4,936,873	3,741,182	4,029,244
	Midwest	9,038,979	8,160,912	7,942,963	5,704,771	5,152,973
	South	13,197,169	12,305,670	12,194,808	10,302,269	9,250,414
	West	6,255,617	5,739,464	5,553,200	3,949,130	3,497,469
dual users	Northeast	620,409	468,171	421,850	404,558	618,139
	Midwest	956,052	859,323	923,093	1,081,676	1,170,175
	South	1,666,177	1,445,659	1,276,672	1,900,223	2,113,033
	West	965,933	636,237	694,194	778,096	895,569
exclusive NVP	All	1,287,030	1,607,448	2,424,175	9,233,782	8,960,790
quitters	Northeast	2,636,428	2,437,112	1,724,364	2,014,465	2,120,397
	Midwest	3,483,994	3,395,944	3,156,303	3,068,646	3,114,284
	South	5,020,674	4,703,122	4,563,909	5,403,957	5,139,652
	West	3,450,164	3,279,153	2,412,810	3,042,929	3,171,442

† Statistical regions of the US Census Bureau

User group	Ethnicity	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	white	25,836,656	23,573,875			
	non-white	8,063,713	7,147,046			
	white non-hispanic			21,175,160	16,075,990	14,882,132
	white hispanic			1,911,070	1,691,013	1,358,553
	black+other			7,541,614	5,930,349	5,689,416
dual users	white	3,508,323	2,846,521			
	non-white	700,248	562,869			
	white non-hispanic			2,461,921	3,042,478	3,607,989
	other			853,890	1,122,075	1,188,927
exclusive NVP	all	1,287,030	1,607,448	2,424,175	9,233,782	8,960,790
quitters	white	11,710,149	11,173,436			
	non-white	2,881,111	2,641,895			
	white non-hispanic			882,3871	9,692,648	9,521,147
	other			3,033,515	3,837,349	4,024,627

User group	Education	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	low	18,201,791	16,410,569	16,512,472	13,617,019	12,040,608
	moderate	11,243,225	10,081,105	9,977,445	7,076,913	6,840,004
	high	4,455,353	4,229,247	4,137,927	3,003,420	3,049,489
dual users	low	1,972,896	1,754,468		2,318,198	2,440,022
	moderate	1,588,640	1,079,841	3,315,811	1,426,635	1,773,916
	high	647,035	575,081		419,720	582,978
exclusive NVP	low			876,938	3,883,339	4,170,298
	moderate	1,287,030	1,607,448	1,106,642	3,527,177	2,865,732
	high			440,595	1,823,266	1,924,760
quitters	low	5,788,777	5,668,481	4,899,488	5,893 992	6,239,134
	moderate	5,504,342	4,897,301	4,244,744	4,628,303	4,401,991
	high	3,298,141	3,249,549	2,713,154	3,007,702	2,904,650

Time since quit	Use of NVP	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
<1 year	yes	1,107,024	844,610	943,699	1,079,441	1,019,173
1 to 5 years		1,161,295	1,018,933	1,381,813	3,302,469	3,118,096
<1 year	no	487,4804	4,745,686	2,994,996	1,863,165	1,916,200
1 to 5 years		744,8137	7,206,102	6,536,878	7,284,922	7,492,306

A.2.3 England

The benchmark/calibration figures used to compute the weights for English respondents are given below. The 2015 Opinions and Lifestyle Survey (OLS) in combination with waves 117–122 (Jun–Nov 2016) of the Smoking Toolkit Study (STS) was used for the wave 1 weights (as well as for all longitudinal weights computed for the wave 1 sample successfully recontacted at follow-up waves; i.e., waves 1–2, waves 1–3,

waves 1–4, etc.). The 2017 OLS in combination with and waves 137–142 (Feb–Jul 2017) of the Smoking Toolkit Study was used for the wave 2 cross-sectional weights (as well as for all longitudinal weights computed for the wave 2 sample successfully recontacted at follow-up waves; i.e., waves 2–3, waves 2–4, waves 2–5, etc.). The 2019 Opinions and Lifestyle Survey (OPN) and Waves 161–164 (Feb–Jun 2020) of the Smoking Toolkit Study was used for the wave 3 cross-sectional weights (as well as the waves 3–4 and 3–5 longitudinal weights). The 2021 & 2022 Opinions and Lifestyle Survey (OPN) and Waves 190–194 (Feb–Jun 2022) of the Smoking Toolkit Study was used for the wave 4 cross-sectional weights (as well as the waves 4–5 longitudinal weights). Finally, the 2023 Opinions and Lifestyle Survey (OPN) and Waves 213–217 (Jul–Nov 2024) of the Smoking Toolkit Study were combined with the 2021 & 2022 OPN and Waves 190–194 of the STS to obtain the wave 5 cross-sectional weights. Combining multiple years of the OPN and STS was required because the sample sizes of those benchmark surveys were not large enough on their own. Consequently, combining made the estimated benchmark figures more reliable and stable.

		# Individuals				
User group	Sex	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	male	3,554,691	3,292,815	2,927,410	1,888,048	1,766,880
	female	2,970,292	2,869,473	2,680,265	1,865,582	1,435,282
dual users	male	523,929	617,110	608,813	712,839	664,229
	female	574,427	454,184	482,243	553,974	550,680
exclusive NVP	male	131,240	274,676	261,482	680,199	779,983
	female			225,872	382,474,389	621,937
quitters	male	1,785,259	1,772,990	1,683,775	1,812,124	1,842,634
	female	1,554,727	1,516,559	1,507,627	1,728,506	1,753,684

		# Individuals				
User group	Age	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	[18, 25)	1,186,770	892,806	852,003	1,370,986	191,512
	[25, 40)	2,225,303	2,084,566	1,900,728	775,286	
	[40, 60)	2,076,676	2,280,389	1,799,006	1,493,112	1,379,785
	[60, 100)	1,036,234	904,527	1,055,938	889,533	855,582
dual users	[18, 25)	102,497	141,948	147,984	619,728	184,248
	[25, 40)	358,135	346,892	407,417		364,356
	[40, 60)	426,513	439,950	398,205	450,151	456,503
	[60, 100)	211,211	142,504	137,450	195,936	209,802
exclusive NVP	[18, 40)			134,281	519,015	701,337
	[40, 60)	131,240	274,676	271,825	377,344	495,262
	[60, 100)			81,248	166,312	668,556
quitters	[18, 25)	391,820	252,258	207,126	1,553,400	263,473
	[25, 40)	1,168,175	1,461,545	1,257,433		1,332,606
	[40, 60)	1,259,689	987,815	1,114,881	1,197,179	1,331,685
	[60, 100)	520,302	587,931	611,962	790,050	668,556

User group	Region	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	East-Midlands	444,925	549,850	543,664	326,697	314,141
	Eastern	759,288	659,255	768,209	378,358	336,354
	London	1,153,042	1,114,846	730,461	596,074	565,103
	North-East	360,254	275,222	353,527	185,536	161,296
	North-West	855,000	775,126	753,721	527,190	400,325
	South-East	979,017	757,313	856,241	529,984	440,190
	South-West	562,755	624,052	473,721	418,446	352,312
	West-Midlands	690,446	662,183	532,669	325,262	312,761
	Yorkshire & Humber	720,256	744,441	595,462	466,085	319,681
dual users	East-Midlands	122,907	59,910	68,459	61,118	95,310
	Eastern	124,848	100,222	101,381	109,902	111,340
	London	98,227	172,863	192,643	192,498	164,786
	North-East	57,281	113,209	86,889	56,911	43,456
	North-West	198,460	88,530	119,231	278,213	192,604
	South-East	120,737	164,949	165,822	185,019	204,204
	South-West	143,277	106,589	123,403	161,340	128,678
	West-Midlands	89,558	100,120	191,784	108,102	176,623
	Yorkshire & Humber	143,061	164,902	41,444	113,710	97,906
exclusive NVP	All	131,240	274,676	487,354	1,062,672	1,401,921
quitters	East-Midlands	286,551	217,673	331,553	258,228	305,963
	Eastern	347,402	319,133	384,529	369,160	364,706
	London	508,325	599,116	509,361	507,453	507,555
	North	599,351	603,113	794,644	574,441	623,265
	South-East	536,668	428,176	381,273	164,368	610,122
	South-West	443,747	346,697	286,417	52,712	385,470
	West-Midlands	250,157	422,847	270,791	440,532	428,695
	Yorkshire & Humber	367,785	352,794	232,834	382,460	370,542

User group	Education	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
cigarette only	low	1,157,268	1,163,421	386,513	715,129	579,314
	moderate	4,402,363	3,805,644	4,645,289	2,424,168	2,047,469
	high	965,352	1,193,223	575,873	614,334	575,380
dual users	low	267,078	141,002	9,608	167,396	146,050
	moderate	691,568	748,559	951,282	916,582	855,792
	high	139,710	181,733	130,166	182,836	213,066
exclusive NVP	low			11,896	97,306	128,649
	moderate	131,240	274,676	371,088	744,214	948,533
	high			104,370	221,154	298,287
quitters	low	505,615	361,278	135,555	521,156	408,722
	moderate	2,199,820	2,092,855	2,605,235	2,308,830	2,482,367
	high	634,551	835,416	450,612	710,645	705,230

Time since quit	Use of NVP	# Individuals Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
< 1 year	yes	434,531	1,065,713	213,605	466,115	419,482
1–5 years		618,021		589,528	862,173	693,070
< 1 year	no	2,287,434	2,223,836	2,388,269	729,085	1,002,300
1–5 years					1,483,258	1,481,468

A.2.4 Australia

The benchmark/calibration figures used to compute the weights for Australian respondents are given below. The 2016 National Drug Strategy Household Survey (NDSHS) was used for both waves 1 and 2 (as well as for all longitudinal weights computed for the waves 1 and 2 samples successfully recontacted at follow-up waves; waves 1–2, waves 1–3, waves 2–3, waves 1–4, waves 2–4, etc.). The 2019 National Drug Strategy Household Survey (NDSHS) was used for both the waves 3 and 4 cross-sectional weights (as well as the waves 3–4 and waves 3–5 longitudinal weights). Lastly, the 2022–23 National Drug Strategy Household Survey (NDSHS) was used for the wave 5 cross-sectional weights.

As mentioned at the end of section 3.1.1, the 2019 NDSHS provided little information to properly estimate the number of HTP users at waves 3 and 4. Hence, some (in particular exclusive NVP) of the benchmark figures (in the middle column of the tables below) are likely to somewhat underestimate the true values. This also applies to the 2022–23 NDSHS.

User group	Sex	# Individuals Waves 1 & 2	Wave 3 & 4	Wave 5
cigarette only	male	1,381,745	1,187,217	846,629
	female	1,086,016	983,536	734,835
dual users	male	58,642	106,750	200,128
	female	34,965	76,450	129,768
exclusive NVP	male	37,063	31,862	122,204
	female		9,911	92,210
quitters	male	489,234	676,259	1,116,285
	female	421,955	527,024	909,937

User group	Age	# Individuals Waves 1 & 2	Wave 3 & 4	Wave 5
cigarette only	[18, 40)	1,192,953	971,546	465,074
	[40, 60)	952,705		
	[40, 55)		696,517	557,860
	[60, 100)	322,103		
	[55, 100)		502,690	558,531
dual users	[18, 40)	52,694	96,956	227,316
	[40, 60)	33,457		
	[40, 55)		59,829	71,810
	[60, 100)	7,456		
	[55, 100)		26,415	30,770
exclusive NVP	[18, 100)	37,063	41,773	214,414
quitters	[18,40)	482,312		
	[18,55)		952,073	1,233,677
	[40, 60)	331,019		
	[60, 100)	97,858		
	[55, 100)		251,210	792,545

User group	State(s)	# Individuals Waves 1 & 2	Wave 3 & 4	Wave 5
cigarette only	NSW+ACT	782,945	661,142	434,759
	VIC	610,502	566,110	381,431
	QLD	573,875	504,016	408,086
	SA	151,126	210,840	127,730
	WA+TAS+NT	349,313	228,645	229,459
dual users	NSW+ACT	25,606	47,920	105,950
	VIC	20,972	38,230	81,726
	QLD	18,300	54,520	73,425
	SA	5,921	42,530	22,416
	WA+TAS+NT	22,808		46,379
exclusive NVP	All	37,063	41,773	214,414
quitters	NSW+ACT	301,084	392,063	710,390
	VIC	224,038	315,659	464,526
	QLD	175,240	254,932	435,472
	SA	60,840	115,568	131,765
	WA+TAS+NT	149,987	125,061	284,068

User group	Education	# Individuals Waves 1 & 2	Wave 3 & 4	Wave 5
cigarette only	Cert I/II, year 12 or less	1,075,843	968,028	1,527,161
	Cert III/IV	701,837	649,225	1,148,274
	University	690,081	553,500	466,359
dual users	Cert I/II, year 12 or less	27,438		273,545
	Cert III/IV	36,028	183,200	236,127
	University	30,141		100,571
exclusive NVP	All	37,063	41,773	829,564
quitters	Cert I/II, year 12 or less	347,887	725,715	501,551
	Cert III/IV	255,100		494,837
	University	308,202	477,568	265,468

Time since quit	Use of NVP	# Individuals Waves 1 & 2	Wave 3 & 4	Wave 5
< 1 year	yes	22,993	111,423	312,568
1–2 years		9,514		
< 1 year	no	647,857	1,091,860	1,713,654
1 to 2 years		230,825		

A.3 Raking algorithm

This section details the raking algorithm used to calibrate the weights to the benchmark figures of section A.2. The description is for the Canadian wave 1 cross-sectional weights (see variable `kWTS100v` in section 1.1.2, and calibration figures of section A.2.1), but the exact same logic applies to waves 2–5 and to the other 3 countries.

Step 1: Let $w_i^{(0)}$ be the start weight of the i^{th} respondent. If this is the very first iteration of the raking algorithm, then $w_i^{(0)} = 1$ for all respondents or (starting at wave 4) is the weight computed using the method of Chen, Li & Wu (2020); otherwise, $w_i^{(0)} = w_i^{(5)}$, as computed in step 5 below.

The $w_i^{(0)}$ weights are calibrated to the user groups × sex benchmark figures of the first table of section A.2.1. To this end, the respondents were divided into the 8 user groups/sex cells of that table. For respondents in cell k ($k = 1, \ldots, 8$), this calibration/post-stratification adjustment consists in multiplying their $w_i^{(0)}$ weights by $\widehat{N}_k / t_k$ to produce calibrated $w_i^{(1)}$ weights. These $w_i^{(1)}$ weights are such that their sum over all respondents in cell k is equal to $\widehat{N}_k$, the estimated number of individuals in that cell (as obtained from the 2015 CTADS Survey). Let k be the cell to which the i^{th} respondent belongs to, the $w_i^{(1)}$ weight of that respondent is given by

$$w_i^{(1)} = w_i^{(0)} \times \frac{\widehat{N}_k}{t_k} = w_i^{(0)} \times \frac{\widehat{N}_k}{\sum_{i \in C_k} w_i^{(0)}}$$

where $\widehat{N}_1, \ldots, \widehat{N}_8$ are given in column 3 of the first table of section A.2.1 and C_k is the set of all respondents in cell k.

Step 2: Using the same post-stratification technique described in step 1, the $w_i^{(1)}$ weights were then calibrated to the user groups × age groups benchmark figures of the second table of section A.2.1. The respondents were divided into the 11 user groups/age groups cells of that table. The $w_i^{(1)}$ weights of respondents in cell ℓ ($\ell = 1, \ldots, 11$) were then multiply by a factor, $\widehat{N}_\ell^{(2)} / t_\ell^{(2)}$, to produce calibrated $w_i^{(2)}$ weights. Let ℓ be the cell to which the i^{th} respondent belongs to, the $w_i^{(2)}$ weight of that respondent is given by

$$w_i^{(2)} = w_i^{(1)} \times \frac{\widehat{N}_\ell^{(2)}}{t_\ell^{(2)}} = w_i^{(1)} \times \frac{\widehat{N}_\ell^{(2)}}{\sum_{i \in C_\ell^{(2)}} w_i^{(1)}}$$

where $\widehat{N}_1^{(2)}, \ldots, \widehat{N}_{11}^{(2)}$ are given in column 3 of the second table of section A.2.1 and $C_\ell^{(2)}$ is the set of all respondents in cell ℓ. It is important to recognize that this second calibration partially destroys the calibration done in step 1; in other words, we no longer necessary have that

$$\sum_{i \in C_k} w_i^{(2)} = \widehat{N}_k \quad \text{for } k = 1, \ldots, 8$$

where C_k and $\widehat{N}_k$ were defined in step 1 above. Because of this, step 1 will need to be repeated (most likely multiple times) after calibrating to the other benchmark figures of section A.2.1; see step 6 below.

Step 3: The $w_i^{(2)}$ weights were then calibrated to the 16 user groups $\times$ region benchmark figures of the third table of section A.2.1. This was done using the same post-stratification technique as detailed above, and yielded the $w_i^{(3)}$ weights. As in step 2, this third calibration partially destroys the calibration done in steps 1 and 2, and those two steps will need to be repeated; see step 6 below.

Step 4: The $w_i^{(3)}$ weights were then calibrated to the 7 user groups $\times$ language benchmark figures of the fourth table of section A.2.1. This was done using the same post-stratification technique as detailed above, and yielded the $w_i^{(4)}$ weights, which partially destroys the calibration done in steps 1–3.

Step 5: For quitters (i.e., group iv as described in section 1.1.2), the $w_i^{(4)}$ weights were further calibrated to the 4 quit subgroups of the last table of section A.2.1. This was done using the same post-stratification technique as detailed above, and yielded the $w_i^{(5)}$ weights. For the other 3 groups (i.e., cigarette only users, dual users, exclusive NVP users), $w_i^{(5)} = w_i^{(4)}$.

Step 6: Repeat steps 1–5 until convergence; that is until,

$$\sum_{i \in C_k} w_i^{(5)} = \widehat{N}_k \quad \text{for } k = 1, \ldots, 8 \qquad \sum_{i \in C_\ell^{(2)}} w_i^{(5)} = \widehat{N}_\ell^{(2)} \quad \text{for } \ell = 1, \ldots, 11$$

and likewise for the calibrations on: user groups $\times$ region (step 3), user groups $\times$ language (step 4), and the 4 quit subgroups (step 5). In other words, repeating steps 1–5 until convergence ensures that the weights are calibrated to all the benchmark figures of section A.2.1.

Note that in the US and Australia, weight trimming (and redistribution) was done at some of the 5 steps mentioned above. This was done to prevent extreme weight variation arising from a few respondents having very large sampling weights.

A.4 Method of Chen, Li & Wu

In this section we briefly explain the method of Chen, Li & Wu (2020). Though their method can be used on its own to compute/produce sampling weights, we opted to instead combine it with our usual raking procedure. This was done to ensure that the weights would be more consistent with those from prior waves as well as with the weights for the ITC England sample, where the benchmark data available to us does not allow us to use the method of Chen, Li & Wu (2020). Lastly, such an approach of combining the method of Chen, Li & Wu (2020) with the raking algorithm, made it simple to trim large weights; something that was highly desirable in our case.

As briefly mentioned in section 4.1.2, their method consists in fitting a logistic regression model. However, this is a slightly different regression model as it utilizes data from two sources/surveys instead of a single dataset. In our case, those two sources/surveys are: *i)* our ITC sample (labelled S_{ITC} in the formulas below) and *ii)* the benchmark survey (labelled S_{BM} in the formulas below); i.e., the NHIS for the US, the CCHS for Canada and the NDSHS for Australia.

Via that binary logistic regression model, the probability of selection or propensity score of the i^{th} individual in our ITC sample S_{ITC} is modelled as a function of a vector of covariates x_i and corresponding vector of parameters θ; i.e., $P(i \in S_{\text{ITC}}) = \pi_i(x_i, \theta)$. In our case, the vector of covariates x_i consists of the following: user group, age, sex, educ, region, ethnic (US only), language (Canada only), user group

× gender, user group × age, user group × education, user group × ethnicity (US only) and user group × language (Canada only). Next, $\hat{\theta}$, the estimator of θ, is obtained by maximizing the following pseudo log-likelihood function:

$$\ell(\theta) = \sum_{i \in S_{\text{ITC}}} x_i^\intercal \theta - \sum_{j \in S_{\text{BM}}} w_j \log\left(1 + \exp(x_j^\intercal \theta)\right)$$

where w_j is the sampling weight of the j^{th} individual in the benchmark survey S_{BM}, and $^\intercal$ simply denotes the transpose. Finally, the start weight of the i^{th} respondent (i.e., the one used in step 1 of our raking algorithm in section A.3) is given by:

$$w_i^{(0)} = \frac{1}{\pi_i(x_i, \hat{\theta})} = \frac{1 + \exp(x_j^\intercal \hat{\theta})}{\exp(x_j^\intercal \hat{\theta})}$$