# ITC Japan Survey Wave 1 and 1.5 Technical Report

1 April 2020

**ITC Japan Research Team**
Dr. Itsuro Yoshimi[1], Principal Investigator
Dr. Yumiko Mochizuki[2], Principal Investigator
Dr. Takahiro Tabuchi[3], Collaborator

**ITC Japan International Research Team**
Professor Geoffrey T. Fong[4] , Principal Investigator, Canada
Professor Mary E. Thompson[5], Co-Investigator, Canada
Dr. Anne C.K. Quah[4], Co-investigator, Canada
Dr. Christian Boudreau[5], Co-investigator, Canada
Ruth Loewen[4], Survey Management Team, Canada
Michelle McAvoy[4], Survey Management Team, Canada
Janine Ouimet[5], Project Manager, Canada
Mi Yan[4], Data Analyst, Canada

[1] Center for Cancer Control and Information Services, National Cancer Center, Japan
[2] Japan Cancer Society
[3] Cancer Control Center, Osaka International Cancer Institute.
[4] Department of Psychology, University of Waterloo
[5] Department of Statistics and Actuarial Science, University of Waterloo

**Suggested Citation:** ITC Project. (2020, April). *ITC Japan Wave 1 (2020) Technical Report.* University of Waterloo, Waterloo, Ontario, Canada, and the Japan National Cancer Center, Tokyo, Japan.

# 1 TABLE OF CONTENTS

**1 SUMMARY AND OVERVIEW OF THE PROJECT**

**1.1 Introduction**

The International Tobacco Control Policy Evaluation Project (ITC Project) was established in 2002 to monitor and evaluate key health policies implemented in countries that are signatories to the Framework Convention on Tobacco Control (FCTC)—the first-ever international public health treaty—that was adopted in May 2003 by all 192 member states of the World Health Organization. Over the past two decades, the ITC Project has provided invaluable data to inform governments and other stakeholders on whether public health policies designed to reduce the health, economic, and societal costs of tobacco use throughout the world, are effective. The ITC Project conducts longitudinal surveys in representative cohorts in over 29 countries including Japan as the 29th ITC country. The ITC Project team from the University of Waterloo in Canada, partnered with the Japan National Cancer Centre and the Japan Cancer Society to develop and field the ITC Japan Survey.

The ITC Japan Project officially commenced in September 2017 with planning and survey development, and the fieldwork was conducted from February 3 to March 2, 2018. The ITC Japan Survey was designed to collect data from Japanese smokers, heated tobacco product (HTP)[1] users, and non-users regarding their knowledge, attitudes, beliefs, perceptions, behaviours, and use patterns associated with cigarettes and heated tobacco products. Japan has the most important market for HTPs, and thus, patterns of use and attitudes from Japan may start to appear in other countries in which these products have been or will be released. One objective of the ITC Japan Survey is to provide evidence to inform tobacco control policies related to HTP products and their use. Although HTP products are marketed as reduced-harm tobacco products, the effect of HTP use on tobacco cessation, uptake, and/or sustained use is not yet understood.

**1.2 Main Objectives and research questions**

The objectives of the Wave 1 of the ITC Japan Survey (JP1) were:
1) To describe how HTP and cigarette use differs among smokers and HTP-only users (some of whom may also be recent ex-smokers) both over time and between important subgroups (e.g., age, gender, income, those planning to quit cigarettes, nicotine dependence level);

2) To examine patterns of tobacco-use behaviour and opinions associated with smoking among adults aged 20 and older in Japan;

3) To examine the effect of a comprehensive smoke-free law on both cigarette smoking and HTP use overall and in public places as well as private places (e.g., homes, cars);

4) To examine the impact of tobacco control policies (other than smoke-free) in Japan.

---

[1] **Note on terminology**: The term "heated tobacco product" or "HTP" is used in the text throughout this report and is the term ITC uses to describe such products as IQOS, glo, and PloomTECH. The term "heat-not-burn product" or "HNB" is used in figures/tables and SAS code in this report, because "HNB" was the English-language term used at the time of JP1 Survey development and fieldwork. Regardless of the English terms, the Japanese term for HTP was used the fielded Japanese-language survey and has remained consistent over time.

**1.3 Overview of project**

- The ITC JP1 Survey was a web-administered survey of behaviours and attitudes related to tobacco and nicotine use among a sample of n=4,615 adult residents of Japan, comprising the following user-type subgroups: 3,288 cigarette smokers, 549cigarette/HTP dual users, 164 HTP-only users, and 614 non-users. The user type definitions are provided in Table 2 in Section 3.1.  The survey took about 27 minutes (median) to complete.   Technical details of the web survey are provided in Appendix 3.

- All sampling and fieldwork was conducted by the survey firm, Rakuten Insight.  The sampling frame was an existing Rakuten Insight panel that was nationally representative of Japanese cigarette smokers, non-users, and HTP users, (see Section 3.1.2) and further quotas based on region of residence, sex, and age, were applied to ensure the final sample was proportional to stratum sizes based on Japan census data.

- The JP1 Survey was conducted from February 3 to March 2, 2018.  Follow up surveys are planned contingent upon funding.

- Rakuten Insight has modified their standard procedures to maximize retention in the panel and the ITC JP study across waves by using a customized ITC-branded background for the survey and survey communications, as well as offering bonus incentives for respondents who complete follow up waves of the survey.


**2 SURVEY MEASURES AND PROGRAMMING**

**2.1 Survey development**
The survey development process comprises four main phases:
  1) determining survey content,
  2) operationalization of survey content,
  3) translation, and
  4) translation review and checking.

**2.1.1  ITC Survey Development process – content and operationalization**
- During Phase 1 of the survey development process, the research investigators, project management team, and the survey management team determined which topics were most important to include in the survey, and then developed the detailed survey questions necessary to measure relevant constructs using the existing framework of the ITC database of questions.  Questions were adapted to the Japan context and new questions were designed as necessary. The resultant draft survey was then sent to the ITC Survey Management Group (SMG) for operationalization of the survey (Phase 2).

- In Phase 2, the operationalization of survey development, involved comprehensively and iteratively reviewing, and revising the survey to ensure that routing, question wording, response options, and all other survey elements are refined and cross-referenced for consistency, clarity, and accuracy.  At the conclusion of Phase 2, the final draft of the survey was generated by SMG and sent to Rakuten Insight for programming and testing.

- During the period when the survey firm programmed and tested the survey, additional revisions were made in consultation between Rakuten Insight and SMG, until a fieldwork version of the survey was achieved.  The fieldwork version of the survey was sent to SMG by

the firm and is retained in the SMG database.  The updated last version of the survey in the database was later used to cross reference with the data set.

### 2.1.2  ITC Survey Development process – translation and review/verification
- The team developed the JP1 survey content and specifications in English initially.  The final English JP1 survey was then translated into Japanese by Rakuten Insight per specifications provided by the research team.

- After the initial translation from English to Japanese had been completed and checked internally by the Rakuten Insight translator(s), the Japanese translation of JP1 was then checked by the Japan team, and issues were identified, discussed, and resolved to confirm the Japanese translation met the research team's standards for the highest possible degree of accuracy.

### 2.1.3 Survey development process – timeline

**Table 1: Summary of the JP1 survey development timeline.**

| Task | Start date | End date |
| --- | --- | --- |
| Survey content determined | April 10, 2017 | September 18, 2017 |
| Operationalization | August 15, 2017 | September 29, 2017 |
| Translation (includes discussion) | November 13, 2017 | November 22, 2017 |
| Translation review and checking | November 23, 2017 | January 8, 2018 |

### 2.2  Survey content

The JP1 survey content was developed to assess the research objectives described in Section 1.2 as well as measure other constructs necessary to meet the survey objectives.  These include demographic, social and psychological factors relevant to models of behaviour change, as well as content to meet logistical requirements for the survey.

The specific JP1 content included the following:
- Information about the survey, time commitment, contact information for ethical concerns or survey-related concerns, and an explicit consent screen.

- Screening section that assesses age, gender, region of residence, smoking status, heat-not-burn product use status

- Heat-not-burn product questions: duration of use, dependence, brand choice, warning labels, purchase, environmental exposure, advertising/promotion, and perceived risk.

- Cigarette questions:  branch choice, perceptions of light/mild, dependence, quitting attempts and aids used, knowledge of health effects of tobacco, warning labels, smoke free places, advertising and promotion,  purchase, beliefs about quitting, psychosocial beliefs, regulation, perceived risk;

- Other questions: anti-tobacco campaigns, e-cigarette use, moderators (i.e., factors not asked in any of the previous categories that are important in models of behaviour change and/or policy evaluation such as time perspective, stress, co-morbidities).

- Demographic questions (e.g., age, gender, education, income, socio-economic status).

## 2.3 Survey programming and testing

- Rakuten Insight used the JP1 specifications (provided in Microsoft Word format) to program the JP1 English Survey using Confirmit Computer-assisted Web Interview (CAWI) Software.

- Rakuten Insight worked closely with the research team to test the English survey CAWI program and survey quotas. The research team provided signoff on the JP1 English Survey program that met pre-determined standards prior to beginning data collection.

- After receiving signoff on the Japanese translation, Rakuten overlaid the Japanese into the CAWI program (originally programmed in English).

- The Japanese translation overlap into the English CAWI program was then systematically tested to ensure the program was functioning as intended and free of errors.


## 3 STUDY SAMPLE

### 3.1 Overview of Wave 1 ITC Japan Survey sample and quotas

The target sample size for the JP1 Survey was 4,250 Japanese adult cigarette smokers, HTP-only users, dual users of cigarette and HTP, and non-users. Table 2 below provides the sub-sample definitions, targets, and achieved sample.

**Table 2. ITC Japan Wave 1 Survey subsample quotas and definition of the user types*.**

| Subsample (quota) group | Definition | Target n | Valid n | Speeders* |
|---|---|---|---|---|
| 1) Current smokers | Smokes cigarettes at least monthly AND (uses HTP not at all or less than weekly). | 3, 000 | 3,288 | 169 |
| 2) HTP only users | (Smokes cigarettes not at all or less than monthly) AND uses HTP at least weekly | 200 | 164 | 9 |
| 3) Cigarette/HTP dual users | Smokes cigarettes at least monthly AND uses HTP at least weekly | 450 | 549 | 25 |
| 4) Non-users | (Smokes cigarettes not at all or less than monthly) AND (uses HTP not at all or less than weekly) | 600* | 614 | 26 |
| **Total** | | **4,250** | **4,615** | **229** |

*Speeders are invalid complete survey records that were not included in the final data set. See Section 5.3 for details.

### 3.1.1 Inclusion/exclusion criteria and quotas
- Rakuten Insight used demographic profile information about their existing panelists to inform which panelists would be invited to the survey.

- Once invited to the survey, the panelists first completed screening to ensure they met the following inclusion criteria:
  - Participants were adults aged 20 years or older;
  - Participants met the definition for one of the four user types specified in Table 2;

- The quota for the panelist's specifications (i.e., user type, region of residence, age, gender) was still open.

  The study exclusion criteria were:
  - Those younger than 20 years old;
  - Corresponding quota to panelist's specifications are full.
  - Note: Following data collection, any identified 'speeders' were removed from the dataset, as described in Section 5.3.

- The study sample was allocated proportionally to stratum sizes based on census data.  The research team established quotas for the JP1 sample based on age, gender, region, and degree of urbanization.  *Refer to Appendix 1: Allocation (per stratum) of JP1 web panel sample.*

### 3.1.2 Description of sampling frame
- The sampling frame in the JP1 Survey was Rakuten Insight's Japan web panel. All of the subsample groups were recruited from Rakuten Insight's Japan panel(s).
- Rakuten Insight provided the following description of their panel(s):  The JP1 survey was conducted with Rakuten's proprietary online panel in Japan. The online panel is actively managed in-house with a dedicated panel management team in Tokyo, and utilized for market research purposes only. Recruitment for the panel is conducted on a daily basis, tapping into users of Rakuten services (e.g. e-commerce, credit cards, insurance, mobile services, etc.) who collectively comprise of roughly 80% of the Japanese population, as well as other online resources such as affiliates, email and banner recruits in order to maintain a panel as consistent as possible with the general population. Panelists are pre-profiled with a series of questions which in turn can be used as pre-targeting variables (e.g. smoking, HTP usage, etc.). Panelists receive email invitations and also have the option of logging into their proprietary panel site to access the survey they are invited to participate in. Details available at: https://insight.rakuten.com/.

## 4 RECRUITMENT AND INTERVIEW PROCEDURES

## 4.1 Contact and recruitment procedures

### 4.1.1 Recruitment strategy
1) During the initial phase of fieldwork invitations only targeted panelists who likely would meet the three tobacco user definitions (Subsample Groups 1-3 in Table 2.  These invitations were termed 'Phase 1 Invitations'. Some of the panelists who were recruited as potential tobacco users actually met the definition of a non-user and thus, were placed in the non-user quota.

2) After the tobacco users quotas were filled, Rakuten Insight invited non-user panelists to the survey, to fill the remaining open positions in the non-user quota (Subsample Group 4 in Table 2).  These invitations were called 'Phase 2 Invitations'.  The sample and subsample categories are provided in Table 2.

### 4.1.2 Invitations and reminders
- All communications with panelists were administered by Rakuten Insight.  All communications were in Japanese.

- Rakuten Insight invited panel members to the JP1 Survey by sending them a standard email invitation that informed the panelists of the survey length and that they would receive the standard incentive for a 35-minute Rakuten Insight survey.

- Per standard procedures, Rakuten Insight sent one email invitation and up to two reminders to panelists who had been pre-identified as being potentially eligible for the JP1 Survey. Once the quotas were achieved, the web survey was closed. Panelists were able to ignore the emails, or contact Rakuten Insight to refuse the study or unsubscribe from the panel at any time.

## 4.2 Fieldwork timeline

- The JP1 Survey was conducted from February 3- March 2, 2018.

## 4.3  The survey experience and interview duration

The JP1 survey was designed to have the look and feel of a typical Rakuten Insight survey, with some branding to identify the survey as an ITC survey.

- The Confirmit software automatically rendered the on-screen formatting to adapt to the respondent's device type (desktop/tablet vs. mobile device) so that text and visual elements would be appropriately placed on the screen to ensure an optimal survey-taking experience.

- The JP1 Survey began with a screening section that assessed panelists' eligibility (based on user type and possibly region/gender/age-based quotas) and determined which survey questions (related to user type) would be asked throughout the survey. Thus, panelists will experience a tailored survey within a single programmed instrument, relevant to each respondent's current tobacco use pattern.

- Consent screens provided information about the survey, time commitment, contact information for ethical concerns or survey-related concerns, and then the panelist was required to provide consent to complete the survey.

- Respondents were able to navigate back to previous questions to change a response.

- Respondents were be able to stop the survey and login to finish at a later time without losing any data.

- Questions were primarily multiple choice format and included one question per page.  Some questions had a check list or grid format, which was modified by the survey software depending on the device type that the respondent used (i.e., desktop vs. a mobile phone).

- There were 22 questions with the 'Other-specify' format, require respondents type open text responses.

- The survey contained encouragement screens.

- Respondents were required to submit their completed survey in order for their survey record to be considered 'complete'.

- Item non-response was acceptable, provided that: the majority of questions were answered, 'essential questions' used for eligibility were answered, and the panelist has submitted their survey.

- The median length of the survey interview was 27 minutes for the valid complete records.

**Table 3. Median survey interview length (minutes)**

| Sample description | n | Median survey interview length (minutes) |
|---|---|---|
| Final valid completes* | 4,615 | 27.37 |
| Completes with speeders | 4,958 | 26.57 |

- 45 inexperienced smokers, 69 respondents who currently smoke daily but have smoked 100+ cigarettes, or who misclassify the HTP and e-cigarettes were excluded from the final valid completes, in addition to 229 speeders.

## 4.4 Assigning disposition codes
- Disposition codes were used to track the outcomes of survey respondents.

- Temporary Disposition Codes were applied to respondents who did not complete the survey within one session.

- Final dispositions codes were assigned to each record (see Section 7 Disposition Codes).

- Three types of disposition codes were used in the study: 1) disposition codes programmed into the survey script, 2) disposition codes entered by the survey firm, and 3) dispositions derived at the end of fieldwork (see Section 7 Disposition Codes).

- Each completed survey record was further sub-coded as being completed on a desktop/tablet device vs. a mobile device vs. being undefined (not possible to classify).

## 4.5 Study incentives
- Panelists were given standard number of points upon completion of the survey, plus bonus points the equivalent to $3 USD.

- Upon submitting a completed survey, the Wave 1 Survey participants were shown a web screen with the information stating that if they stayed in the panel and completed the Wave 2 Survey about 12-18 months later, they would be provided with bonus points equivalent to $10 USD (for cigarette smokers or non-users) or $25 USD (for HTP users).

## 5 QUALITY CONTROL

## 5.1 Fieldwork monitoring and progress reports
- At the beginning of fieldwork, the initial sample invitations were released carefully at deliberate intervals and survey activity was closely monitored to ensure that all aspects were working as intended.  This method is termed a 'soft launch' and occurred from February 3–13, 2018.

- The 'soft launch' data were systematically reviewed by both Rakuten and research team.  No major issues of concern were determined.

- Throughout fieldwork Rakuten Insight closely monitored survey activity and ensured a smooth implementation.

- Rakuten Insight provided the research team with a portal link to view a report monitoring survey activity in live time, and also provided weekly fieldwork reports and an analysis of next steps with respect to the survey recruitment strategy.

## 5.2 Survey completes vs. partial completes

- The definition of a "survey complete" is the survey record for a panelist who started the survey, completed the survey questions, perhaps endorsing "prefer not to answer" for a reasonable proportion of questions, and then chose to 'submit the survey' after the last survey question.

- Survey response data for survey completes were checked using the criteria defined in Section 5.3.  Records that passed the checks were considered valid completes.

- Survey response data for partially completed survey records (defined as records for which the panelist started the survey but did not hit submit at the end of the survey) were not included in the final data set.

## 5.3 Identification and removal of 'satisficers/speeders' from the data set

- Rakuten Insight and ITC's data analysts systematically analysed the submitted survey records to determine which records met the established criteria for a speeder.

- The strategy was to create a group of normal respondents by dropping all poor-quality respondents (those with very low SecperQ and/or high %RDK) and to use this normal group to calculate normal behaviour ranges.  Then speeder behaviour could be defined relative to the normal behaviour (see Table 5).

- Speeder data was removed from the final data set.  Table 3 shows the median interview lengths (minutes) before and after the speeders' records were removed from the dataset.

## 5.3.1 Criteria for assessing poor-quality data

- Two criteria were used to assess poor-quality data:
    1) Seconds per question (SecperQ)
    2) % of responses that were either Refused or Don't Know (%RDK).

- Very extreme values occurred for both of these variables: SecperQ times of less than 1.7 seconds per question, which by published estimates does not allow time for even reading the question, and %RDK responses for more than 70% of the questions completed.
    o Six user groups were created since the respondents are required to answer different series of questions.
        1. Cigarette-only users
        2. HTP-only users
        3. Dual users
        4. Non-users
        5. Short-term quitters using HTP at least weekly (quit within two years)
        6. Short-term quitters using HTP less than weekly (quit within two years)

**5.3.2 Removing poor-quality data to create 'normal' group**

- The initial criteria for creating the normal groups for each user group were high and somewhat arbitrary, as the goal was to remove users who might not have been answering responsibly.

**Table 4: SecperQ and %RDK cut-offs used to determine users ineligible for normal group**

| User Group | SecperQ Cutoff (seconds)[a] | %RDK Cutoff [b] | Users Ineligible for Normal Group (n) |
|---|---|---|---|
| Cig-Only | ≤ 5.700 | ≥ 20.77% | 1042 |
| HTP-only | ≤ 5.440 | ≥ 19.17% | 36 |
| Dual | ≤ 4.795 | ≥ 18.40% | 177 |
| Non-User | ≤ 5.890 | ≥ 22.13% | 188 |
| Short term QT using HTP at least weekly | ≤ 5.650 | ≥ 14.15% | 29 |
| Short term QT using HTP less than weekly | ≤ 6.100 | ≥ 19.44% | 25 |

[a]Users with SecperQ less than or equal to the 20th percentile of their user-group
[b]Users with %RDK greater than the 85th percentile of their user-group

- 1497 out of 4958 respondents were excluded at this stage.

- Frequency distributions by normal user groups were used to create more precise cut-offs for 'poor-quality' responding. 'Faulty points' were assigned based on the normal group percentiles. Respondents were removed from the dataset if they scored too many points.

**5.3.3 Assigning points to user groups, identifying speeders**

- 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 the normal group's 0.5*<u>minimum</u>.

- **3 points** were assigned if the respondent's value fell between the normal group's 0.5*<u>minimum</u> (inclusive) to the 0.75*<u>minimum</u>.

- **2 points** were assigned if the respondent's value fell between the normal group's 0.75*<u>minimum</u> (inclusive) to <u>minimum</u>

- **All other** values would have a point of 0.

Because SecperQ (time per question) was considered the most important of the 2 data-quality criteria, it was weighted more heavily, so the points assigned here were 2, 3 and 5, rather than 1, 2 and 3 as for the %RDK criterion.

**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 ≥ to the normal group's 2*maximum.
>
> - **2 points** was assigned if the respondent's value fell between the normal group's 1.33*maximum (inclusive) to 2*maximum
>
> - **1 point** was assigned if the respondent's value fell between the normal group's maximum to 1.33*maximum
>
> - All other values of %RDK would have a point of 0.

- Points results: The range of possible points was 0 – 8. At the applied cut-off of 4 points, 229 respondents were considered speeders.

**Table 5: Number of speeders and non-speeders by user group**

| User Group | Speeders | Non-speeders | Total |
|---|---|---|---|
| Cig-Only | 169 | 3,288 | 3,457 |
| HTP-Only | 5 | 74 | 79 |
| Dual | 25 | 549 | 574 |
| Non-User | 23 | 536 | 559 |
| Short term QT using HTP at least weekly | 4 | 90 | 94 |
| Short term QT using HTP less than weekly | 3 | 78 | 81 |
| **Total** | **229** | **4615** | **4844*** |

 *45 inexperienced smokers, 69 respondents who currently smoke daily but have smoked 100+ cigarettes, or who misclassify the HTP and e-cigarettes were excluded.

- Comparison between respondents speeders and non-speeders:

> - There were more speeders in the 20-29 (8%) and 30-39 (6.35%) age groups than in the 40-59 (4.88%) and 60+ (1.4%) age groups.
>
> - There were more male speeders (5.11%) than female speeders (3.43%).
>
> - The distribution of speeders across user groups were similar, ranging from 3.70% to 4.86%.

**5.4 Data cleaning and topline frequencies**
- After fieldwork was completed, Rakuten Insight cleaned the data and then transferred the cleaned data to the ITC Project.

- ITC Project analysts completed further data cleaning, weights construction, and conducted initial descriptive analyses, including generating topline frequencies. During this step, the team decided to ask a follow up survey to capture the number of cigarettes used per day. See Section 6 or details**.**

**5.5 Translation review and verification**
- Standard procedures at ITC include validating the translation against the fieldwork-version of the survey. This process is conducted by an independent reviewer fluent in Japanese and English.

- The review process identified three translation issues (variables ET634, CH879, PR101) in which the translation was evaluated to be different enough from the original English question to warrant a revision. The three affected variables were documented and a more accurate translation will be used in future waves.

- For variable ET634, the correctly-translated data were collected in the JP1.5 cohort survey (see Section 6). This variable was included because it is often used in ITC cross-country comparisons.

**5.6 CHERRIES checklist**
- The Checklist for Reporting Results of Internet E-Surveys (CHERRIES) is an assessment tool endorsed by the Journal of Medical Internet Research to promote complete and accurate standard reporting guidelines for authors describing internet-based surveys.

- Appendix 3 provides the CHERRIES checklist for the JP1 Survey.

**6 ITC JAPAN WAVE 1.5 COHORT SURVEY**

**6.1 Purpose and overview of the Japan Wave 1.5 Cohort Survey**
- The ITC Japan Wave 1.5 (JP1.5) Survey was a follow-up to the JP1 Survey.
- The purpose of the JP1.5 Survey was to collect two data fields that were not included in the ITC JP1 Survey.  The fields were **cigarettes consumed per day** (**CPD),** collected via variables FR216, FR226, and FR236 (see Table 4 for survey questions) and ET634, with a revised translation from the JP1 version (Table 4).
- Data were collected from the appropriate respondents who agreed to participate. Eligibility was based on responses to the ITC JP1 Survey.
- The JP1.5 Survey was a Recontact Survey only, designed to obtain data to supplement the JP1 data set.   The cohort was not replenished (i.e., respondents lost to follow up were not replaced by new recruits) at the JP1.5 Survey.
- The ITC team developed the survey questions in consultation with Rakuten Insight.
- Rakuten Insight conducted the programming, translation into Japanese, and all survey fieldwork.
- Ethics clearance was obtained from a University of Waterloo Research Ethics Committee prior to fieldwork initiating.

**Table 6. The ITC Japan Wave 1.5 Survey questions FR216, FR226, FR236 and ET634.**

| Variable name | JP1.5 questions |
|---|---|
| FR216 | **Ask if FR309v@LSD=1.**<br>When you completed our survey in **[February/ March 2018]** you said you were smoking cigarettes every day.<br>    On average, how many cigarettes were you smoking **each day** at the time you completed our survey in **[February/ March 2018]**, including both factory-made and hand-rolled cigarettes?<br>        8888    Refused<br>        9999    Don't Know<br>*(range 1-999)*<br>*This is a very important question. If you don't have an exact answer, do your best to give an estimate.*<br>**Go to ET634.** |
| FR226 | **Ask if FR309v@LSD=2.**<br>When you completed our survey in **[February/ March 2018]** you said you were smoking cigarettes at least once a week.<br>    On average, how many cigarettes were you smoking **each week** at the time you completed our survey in **[February/ March 2018]**, including both factory-made and hand-rolled cigarettes?<br>        8888    Refused<br>        9999    Don't Know<br>*(range 1-999)*<br>*This is a very important question. If you don't have an exact answer, do your best to give an estimate.*<br>**Go to ET634.** |
| FR236 | **Ask if FR309v@LSD=3.**<br>When you completed our survey in **[February/ March 2018]** you said you were smoking cigarettes at least once a month.<br>    On average, how many cigarettes were you smoking **each month** at the time you completed our survey in **[February/ March 2018]**, including both factory-made and hand-rolled cigarettes?<br>        8888    Refused<br>        9999    Don't Know<br>*(range 1-999)*<br>*This is a very important question. If you don't have an exact answer, do your best to give an estimate.* |
| ET634 | **Ask if ET603@LSD=1.**<br>When you completed our survey in **[February/ March 2018]** you said you were employed outside the home and usually worked inside a building.<br>    In the **30 days before** you completed **our tobacco survey on [JP1 survey date]**, had people smoked in indoor areas where you work?<br>        1   Yes<br>        2   No<br>        8   Refused<br>        9   Don't know<br>*This is a very important question. Please do your best to answer.* |

### 6.2 Sample and inclusion/exclusion criteria

- The JP1.5 sample is a subset of respondents who completed the JP1 Survey, but whom we would like to ask 1-2 additional questions to be answered as though it were the time of the JP1 Survey (i.e., February – March 2018).
- The JP1.5 sample is summarized in Table 5, below.
- The inclusion criteria are: adults aged 20 or older, current Rakuten panelists who are eligible to answer one or two of variables FR216, FR226, FR236, or ET634, as determined by previous responses to the JP1 Survey. Question wordings are provided in Table 4.
- Exclusion criteria are: Absence of any inclusion criterion.

**Table 7. The ITC Japan Wave 1.5 Sample as a function of 'cigarettes per day' (CPD, measured by FR216, FR226, FR236) and ET634.**

| Questions respondents will be asked | Number of JP1 respondents eligible for question(s) |
|---|---|
| CPD and ET634 | 2347 (FR216 =2193, FR226 =135, FR236 =19) |
| Only CPD | 1514 (FR216 =1456, FR226 =50, FR236 =8) |
| Only ET634 | 429 |
| **Total** | **4,290** |

### 6.3 Invitations and reminders

- All communications with panelists were administered by Rakuten Insight. All communications were in Japanese.

- Rakuten Insight invited panel members to the JP1 Survey by sending them a standard email invitation that informed the panelists of the survey length and that they would receive the standard incentive for a 1-2-minute Rakuten Insight survey. The email template was similar to that from Wave 1, provided in *Appendix 2*.

- Per standard procedures, Rakuten Insight sent one email invitation and up to two reminders to panelists who had been previously identified as eligible for the JP1 Survey. Once the quotas were achieved, the web survey was closed. Panelists were able to ignore the emails, or contact Rakuten Insight to refuse the study or unsubscribe from the panel at any time.

### 6.4 Fieldwork timeline

- The JP1 Survey was conducted from June 12-20, 2018.

### 6.5 The survey experience and interview duration

- Procedures were identical to those described in *Section 4.3*, with the exception of survey length, which was 1-2 minutes.

### 6.6 Study incentives

- Panelists were given standard number of points for a 1-2 minute Rakuten Insight panel survey. Points were allocated the respondent's account after submitting the completed survey.

- Upon submitting a completed survey, the Wave 1.5 Survey participants were shown a web screen stating that the research team would contact them within the year.

**6.7 Fieldwork monitoring and progress reports**
- At the beginning of fieldwork, the initial sample invitations were released carefully at deliberate intervals and survey activity was closely monitored to ensure that all aspects were working as intended. This method is termed a 'soft launch' and occurred during June 12-13, 2018 (and the full fieldwork then followed during June 14-20, 2018).

- The 'soft launch' data were systematically reviewed by both Rakuten Insight and research team. No major issues of concern were determined.

- Throughout fieldwork Rakuten Insight closely monitored survey activity and ensured a smooth implementation.

- Rakuten Insight provided the research team with a portal link to view a report monitoring survey activity in live time, and also provided regular fieldwork reports.


**6.8 Wave 1.5 survey response and merger with Wave 1 data**

**6.8.1. Data delivery, cleaning, and merger**
- After fieldwork was completed, Rakuten Insight cleaned the data and then transferred the cleaned data to the ITC Project.

- The ITC data analyst further cleaned the data. The Wave 1.5 data were then merged (based on a unique identifier) with the Wave 1 ITC Japan Survey data.

**6.8.2 Survey response**
- Of the eligible sample, 81% completed the Wave 1.5 Japan Survey.

**6.8.3 Imputing missing values for cigarettes consumed per day**
- In the case of missing data/non-response at Wave 1.5 for cigarettes per day, the responses were imputed. Details are provided in Appendix 5.
- Missing data/non-response for ET634 at Wave 1.5 were treated as missing data for variable ET634.

## 7 DISPOSITION CODES

### 7.1 Disposition Codes for Wave 1 ITC Japan Survey

| DMC Code | Type | Description | Comments |
|---|---|---|---|
| **Interview** | | | |
| P-A1* | P | Selected respondent completes the entire survey; maybe skipping or refusing to answer a few questions | *All records coded P-A1 were further subcoded as desktop/tablet vs. mobile device vs. undefined, per Section A4.2. |
| **Eligible, non-interview** | | | |
| P-B19 | E | Respondent completed eligibility questions and was deemed to be eligible, then started to answer the survey but does not complete the survey. | |
| P-B90 | S | Any other reason why interview was not completed, but eligibility was confirmed by respondent | Unlikely to be used, but left in as a precaution |
| **Unknown eligibility, non-interview** | | | |
| P-C11.1 | P | Respondent refuses, can't answer or doesn't know his/her age; thus unknown if he/she is eligible | |
| P-C11.2 | P | Respondent refuses, can't answer or doesn't know his/her gender ; thus unknown if he/she is eligible | |
| P-C11.5 | P | Respondent refuses, can't answer or doesn't know his/her cigarette smoking status; thus unknown if he/she is eligible | |
| P-C11.6 | P | Respondent refuses, can't answer or doesn't know his/her heat-not-burn (HTP) user status; thus unknown if he/she is eligible | |
| P-C13 | P | Respondent refuses at consent; thus unknown if he/she is eligible | |
| P-C70 | S | Invalid email or email bounce back | |
| P-C72 (Formerly P-C75) | S | Respondent never logged into system to start the survey (but there was no email bounce back/invalid) | This is an important disposition code, and we expect that many individuals will fall into this category |
| P-C80 | P | Respondent refuses (or doesn't know) to provide required information to derive their stratum; hence sampling weights cannot be computed and unknown if he/she is eligible | |

| | | | |
|---|---|---|---|
| P-C90 | S | Other reason why unknown eligibility | Unlikely to be used, but left in as a precaution |
| **Not eligible** | | | |
| P-D10 | P | Respondent is out of the target population | For example, respondent does not reside in any of the 47 Prefectures of Japan |
| P-D70 | P | Respondent is too young (i.e., <20 years old) | |
| P-D80.1 | P | Corresponding quota for cigarette smokers only (HN350=1) is full | |
| P-D80.2 | P | Corresponding quota for HTP-only users (HN350=2) is full | |
| P-D80.3 | P | Corresponding quota for dual users (HN350=3) is full | |
| P-D80.4 | P | Corresponding quota for non users (HN350=4) is full | |
| P-D90 | S | Any other reason why respondent is not eligible | Unlikely to be used, but left in as a precaution |

**\*Sub-code for participants coded P-A1: DskTab vs. Mob vs. Undef**

- All records coded P-A1 will further be sub-coded per the definitions below.
    - Desktop/Tablet (DskTab) = Respondent completed survey on desktop computer/tablet defined as a device on which the survey appears large-screen formatting, e.g. grid questions show as grids.
    - Mobile (Mob) = Respondent completed survey on mobile device defined as a smaller-screen device on which the JP1 survey questions are reformatted for a mobile experience, e.g. grids are shown as a list of items that can be scrolled down vertically instead of in a grid that populates horizontally.
    - Undefined (Undef) = Respondent completed survey on device not classified as either a Desktop/Tablet or a Mobile.

Notes:

*1 Type of disposition codes:

P = disposition code programmed into the script

S = Dispositions to be entered by supervisor or other person responsible for fieldwork

E = Dispositions to be derived at the end of fieldwork

HTP = Heat not burn

## 7.2 Disposition Codes for Wave 1.5 ITC Japan Survey

**Recontact Disposition Codes**
**ITC Japan Survey – Wave 1.5**
Last updated by C. Boudreau on May 31, 2018

Notes:

1) Disposition codes are in column B
2) See legend at bottom

| Disposition Codes | | | |
|---|---|---|---|
| DMC Code | Type[1] | Description | Comment |
| **A – Interviewed** | | | |
| C-A1 | P | Selected respondent completes the entire survey; maybe skipping or refusing to answer a few questions | |
| **B – Eligible, non-interview** | | | |
| C-B11.1 | P | Respondent refuses, can't answer or doesn't know his/her age | |
| C-B13 | P | Respondent refuses at consent | |
| C-B15 | P | Age entered by respondent doesn't match age on file for that same respondent | |
| C-B19 | P or E | Respondent consented and started to answer the survey, but did not complete the survey | |
| C-B70 | S | Invalid email or email bounce back | |
| C-B72 | E | Respondent never logged into system to start the survey (but there was no email bounce back/invalid) | This is an important disposition code, and we expect that quite a few individuals will fall into this category |
| C-B90 | S | Any other reason why interview was not completed | Unlikely to be used, but left in as a precaution |
| **D – Not eligible** | | | |
| C-D10 | S | Respondent is out of the target population | For example, respondent has move and now resides in any of the 47 Prefectures of Japan |
| C-D90 | S | Any other reason why respondent is not eligible | Unlikely to be used, but left in as a precaution |

Notes:
[1] Type of disposition codes:

P = disposition code programmed into the script
S = Dispositions to be entered by supervisor or other person responsible for fieldwork
E = Dispositions to be derived at the end of fieldwork

## 8 COOPERATION AND RESPONSE RATES FOR WAVE 1 ITC JAPAN SURVEY*

*Note: Cooperation and response rates were calculated for the Wave 1 ITC Japan Survey only (i.e., rates were not calculated for the Wave 1.5 ITC Japan Survey), because the data collected at Wave 1.5 were merged with the Wave 1 data (Section 6.8).

### Response and cooperation rates – ITC Japan Wave 1
Last updated on Apr 1, 2020 by C. Boudreau

Remark: Because respondents were recruited via Rakuten's web panel, the response and cooperation rates in this spreadsheet are not comparable to those from ITC surveys conducted via phone/RDD or face-to-face

|  | Freq | % |
|---|---|---|
| A – Interviewed | | |
| Total (interviewed)*1 | 4,958 | 17.5% |
| B – Eligible, but not interviewed | | |
| Refusal/breaks off | 164 | 0.6% |
| Other | 0 | 0.0% |
| Total (eligible but not interviewed) | 164 | 0.6% |
| C – Unknown if eligibility (not interviewed) | | |
| Logged into system to start survey (once or more) | 2,724 | 9.6% |
| Estimated number of eligible and quota not full*2 | 1,019 | 3.6% |
| Estimated number of not eligible or quota full*3 | 1,705 | 6.0% |
| Never logged into system to start survey | 11,990 | 42.2% |
| Estimated number of eligible and quota not full*4 | 4,485 | 15.8% |
| Estimated number of not eligible or quota full*4 | 7,505 | 26.4% |
| Total (unknown if eligible) | 14,714 | 51.8% |
| D – Not eligible | | |
| Out of sample | 3 | 0.0% |
| Respondent is not eligible | 3 | 0.0% |
| Quota full | 8,564 | 30.1% |
| Other | 0 | 0.0% |
| Total (not eligible) | 8,570 | 30.2% |
| Total sample with final disposition | 28,406 | 100% |
| Estimated eligibility rate*5 | 99.9% | |
| Estimated proportion for which quota was full*6 | 62.6% | |
| Response rate*7 | 46.7% | |
| Cooperation rate*8 | 96.8% | |

**Notes:**

*1 The total number of completed interviews (i.e., 4,958) includes some respondents that were then removed from the final dataset because they were deemed to be fraudulent/speeder
If those respondents were excluded from the above table, both the response rate and cooperation are would be slightly lower than their current values

*2 Estimated number of respondents that would have been eligible and for which the corresponding quota would not have been full
Formula: row 16 x row 33 x (1 - row 34), rounded to the nearest integer

*3 Formula: row 16 - row 17

*4 Same as *2 and *3

*5 Estimated proportion of individuals that were eligible (i.e., age 20 & older and meet smoking/tobacco use criteria)
Formula: 1 - row 26 / (row 8 + row 13 + row 26)

*6 Estimated proportion of individuals that were terminated because the corresponding quota was full
Formula: row 27 / (row 8 + row 13 + row 27)

*7 The response rate is the proportion of selected/contacted respondents who complete the survey; i.e.,
Formula: row 8 / (row 8 + row 13 + row 17 + row 20)

*8 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 8 / (row 8 + row 11)
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.

## 9 REFERENCES

[1] Tabuchi T, Gallus S, Shinozaki T, Nakaya T, Kunugita N, Colwell B. Heat-not-burn tobacco product use in Japan: its prevalence, predictors and perceived symptoms from exposure to secondhand heat-not-burn tobacco aerosol. Tobacco control. 2018 Jul 1;27(e1):e25-33.

# APPENDIX 1: ALLOCATION (PER STRATUM) OF WAVE 1 ITC JAPAN SURVEY WEB PANEL SAMPLE.

**Quotas for Smokers**

**ITC Japan - Wave 1**

Created by C. Boudreau on Nov 3, 2017

Last updated by C. Boudreau on Nov 10, 2017

Notes:

1) This spreadsheets contains the quotas for the 3000 current smokers*; see other spreadsheets in this Excel file for the other goups.

2) The quotas to be programmed are in orange (see table in cells B13-J29); the other numbers in that table are totals that do not need to be programm

3) Our preferred option is for all quotas to be meet or slightly exceeded; however, we recognize that this might be challenging for some age/gender groups, and are thus somewhat flexible.

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

* In the ITC Japan Survey, a current smoker is defined as someone who smokers cigarettes at least monthly and uses HNB not at all or less than weekly.

| | Quotas | | | | |
| --- | --- | --- | --- | --- | --- |
| | Males (BI240=1) | | | | |
| | 20-29 | 30-39 | 40-59 | 60 & older | |
| | AgeGrp=1 | AgeGrp=2 | AgeGrp=3 | AgeGrp=4 | Total |
| Hokkaido + Tohoku (jpQRegionv=1) | 38 | 66 | 137 | 100 | 341 |
| Kanto (jpQRegionv=2) | 107 | 176 | 338 | 189 | 810 |
| Chubu + Kansai (jpQRegionv=3) | 95 | 156 | 318 | 206 | 775 |
| Chugoku + Shikoku + Kyushu & Okinawa (jpQRegionv=4) | 50 | 86 | 171 | 126 | 433 |
| Total Japan | 290 | 484 | 964 | 621 | 2359 |
| | Females (BI240=2) | | | | |
| | 20-29 | 30-59 | | 60 & older | |
| | AgeGrp=1 | AgeGrp=2 or 3 | | AgeGrp=4 | Total |
| Hokkaido + Tohoku (jpQRegionv=1) | | 60 | | 27 | 87 |
| Kanto (jpQRegionv=2) | | 141 | | 49 | 190 |
| Chubu + Kansai + Chugoku + Shikoku + Kyushu & Okinawa (jpQRegionv=3 | | 213 | | 91 | 304 |
| Total Japan | 60 | 414 | | 167 | 641 |
| | Both | | | | |
| | 20-29 | 30-59 | | 60 & older | Total |
| Total Japan | 350 | 1862 | | 788 | 3000 |

# Quotas for HTP users and dual users

Created by C. Boudreau on Nov 3, 2017

Last updated by C. Boudreau on Nov 10, 2017

Notes:

1) This spreadsheets contains the quotas for the 200 HNB only users* and the 450 dual users*; see other spreadsheets in this Excel file for the other goups.

2) The quotas to be programmed are in orange (see tables in cells B14-G20 and B22-G27); the other numbers in that table are totals that do not need to be programmed.

3) Our preferred option is for all quotas to be meet or slightly exceeded; however, we recognize that this might be challenging for some age/gender groups, and are thus somewhat flexible.

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.

* In the ITC Japan Survey, a dual user is defined as someone who smokes cigarettes at least monthly and uses HNB at least weekly

* In the ITC Japan Survey, an HNB only user is defined as someone uses HNB at least weekly and smokes cigarettes not at all or less than monthly

| | HNB Only[*1] | Population Estimates[*2] | Estimated # of HNB only users | Quotas |
|---|---|---|---|---|
| Males (BI240=1) | 3.28% | 52,690,000 | 1,727,541 | 114 |
| Females (BI240=2) | 2.27% | 56,951,000 | 1,294,341 | 86 |
| Total Japan | | 109,641,000 | 3,021,882 | 200 |

| | | Dual[*3] | Population Estimates[*2] | Estimated # of dual users | Quotas |
|---|---|---|---|---|---|
| Males | | | | | |
| (BI240=1) | 20-29 (AgeGrp=1) | 9.52% | 6,672,000 | 635,429 | 87 |
| | 30 & older (AgeGrp=2-4) | 2.91% | 46,018,000 | 1,341,332 | 185 |
| | Total | | 52,690,000 | 1,976,761 | |
| Females | | | | | |
| (BI240=2) | All | 2.27% | 56,951,000 | 1,294,341 | 178 |
| Total Japan | | | 109,641,000 | 3,271,102 | 450 |

Notes:

*1 Estimated proportion of individuals that are heat-no-burn (HNB) users only

Those estimates are from the HNB incidence check performed by Rakuten in July amongst 901 of their Tokyo panellists

*2 Population estimates/projections as of Oct 1, 2016

Downloaded from e-Stat portal (http://www.e-stat.go.jp/SG1/estat/eStatTopPortalE.do) of the National Statistics Center (http://www.nstac.go.jp/en/)

See spreadsheet Population in this Excel file

*3 Estimated proportion of individuals that are dual users

Those estimates are from the HNB incidence check performed by Rakuten in July amongst 901 of their Tokyo panellists

**Quotas for Non-Users**

**ITC Japan - Wave 1**

Created by C. Boudreau on Nov 3, 2017

Last updated by C. Boudreau on Nov 10, 2017

Notes:

1) This spreadsheets contains the quotas for the 600 non-users; see other spreadsheets in this Excel file for the other goups.

2) The quotas to be programmed are in orange (see table in cells B12-D13); the other numbers in that table are totals that do not need to be programmed.

3) Our preferred option is for all quotas to be meet or slightly exceeded; however, we recognize that this might be challenging for some age/gender groups, and are thus somewhat flexible.

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 |
|---|---|---|
| Males | | |
| (BI240=1) | 20-29 (AgeGrp=1) | 31 |
| | 30-39 (AgeGrp=2) | 32 |
| | 40-59 (AgeGrp=3) | 76 |
| | 60 & older (AgeGrp=4) | 105 |
| | Total | 244 |
| Females | | |
| (BI240=2) | 20-29 (AgeGrp=1) | 40 |
| | 30-39 (AgeGrp=2) | 47 |
| | 40-59 (AgeGrp=3) | 107 |
| | 60 & older (AgeGrp=4) | 162 |
| | Total | 356 |
| Total Japan | | 600 |

**APPENDIX 2: ITC JAPAN EMAIL TEMPLATES**

**Appendix 2.1: Wave 1 ITC Japan SURVEY Email invitation Template**

**Email Subject** Invitation to a Special "3-Year Survey": Users and non-users of tobacco products [PROJECT NUMBER]

**Email Body**:
[NAME],

Thank you very much for participating in the Rakuten surveys.

Warm welcome to the Special "3-Year Survey"!

You are part of a special group of participants across Japan (and in 28 other countries throughout the world) that is invited to participate in this research that will help policymakers and health professionals in forming future policies and programs on tobacco.

The International Tobacco Control (ITC) Japan Survey is a research of people's knowledge, attitudes, and behaviours about tobacco products. Both "users" and "non-users" of tobacco products are needed for this study — a survey that will be conducted about once a year over three years.

This is a continuing research study that involves:
- participating in this 35-minute survey during the next few days, and
- if you complete the survey this time, being invited to participate in follow-up surveys every 12-18 months.

If you qualify and fully complete this survey, it's extremely important that you come back and join our Special "3-Year Survey" again in 12-18 months!

Please join the below survey.

- **Survey URL: [LINK TO SURVEY]**
  - o Answering a survey can only be done once per person.

- Survey length: About 35 minutes. (Actual length may vary depending on contents)
  - o Incentive: [**Rakuten min – max incentive**].

- Closing Date: Once the required number of survey completions is reached.
- Incentive Awarded Date：Max 60 days after completing the survey
  - o Depending on the survey, it may take longer than the above timing.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*
Please take note while answering surveys:
- Please make sure to review each question and answer choice carefully to make sure your answers are accurately selected.

- We suggest the below internet environment: [xx]

- We strictly adhere to the Privacy Policy in the Member Agreement.  Please do not disclose any information you acquired in the survey to any third party, including posts to bulletin boards and blogs.

- Surveys that were emailed more than 2 weeks ago will not appear on your My Page.  Please click the URL to answer the survey.
    - If the URL is cut off based on your email setting, please copy the whole URL and paste it into your browser address bar.
    - You may not enter the survey if the survey is already closed.

- You will not be able to reply to this email.

- Please go to FAQ for any questions.

**Appendix 2.2: Wave 1 ITC SURVEY Email Reminder template #1**

**Email Subject**: Invitation to a Special "3-Year Survey": Users and non-users of tobacco products [PROJECT NUMBER] - Reminder

**Email Body**:
[NAME],

Thank you very much for participating in the Rakuten surveys.

We recently sent you an invitation to the Special "3-Year Survey".

**Your participation is important to us -- we hope you will agree to participate in our study!**

You are part of a special group of participants across Japan (and in 28 other countries throughout the world) that are invited to participate in this research that will help policymakers and health professionals in forming future policies and programs on tobacco.

The International Tobacco Control (ITC) Japan Survey is a research of people's knowledge, attitudes, and behaviours about tobacco products. Both "users" and "non-users" of tobacco products are needed for this study — a survey that will be conducted about once a year over three years.

This is a continuing research study that involves:
- participating in this 35-minute survey during the next few days, and
- if you complete the survey this time, being invited to participate in follow-up surveys every 12-18 months.

If you qualify and fully complete this survey, it's extremely important that you come back and join our Special "3-Year Survey" again in 12-18 months!

We'd appreciate you taking the survey at your earliest convenience by using the web address below.  If you have already completed the survey, thank you and please discard this email.

Please join the below survey.

- **Survey URL:  [LINK TO SURVEY]**
  - Answering a survey can only be done once per person.

- Survey length: About 35 minutes. (Actual length may vary depending on contents)
  - Incentive: [**Rakuten min – max incentive**].

- Closing Date: Once the required number of survey completions is reached.
- Incentive Awarded Date：Max 60 days after completing the survey
  - Depending on the survey, it may take longer than the above timing.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Please take note while answering surveys:

- Please make sure to review each question and answer choice carefully to make sure your answers are accurately selected.

- We suggest the below internet environment: [xx]

- We strictly adhere to the Privacy Policy in the Member Agreement.  Please do not disclose any information you acquired in the survey to any third party, including posts to bulletin boards and blogs.

- Surveys that were emailed more than 2 weeks ago will not appear on your My Page.  Please click the URL to answer the survey.
    - If the URL is cut off based on your email setting, please copy the whole URL and paste it into your browser address bar.
    - You may not enter the survey if the survey is already closed.

- You will not be able to reply to this email.

- Please go to FAQ for any questions.

**Appendix 2.3: Wave 1 ITC Japan SURVEY Email Reminder template #2**

**Email Subject**: Special Survey "3 Year Tobacco Panel" [PROJECT NUMBER] - Reminder

**Email Body**:

[NAME],

Thank you very much for participating in the Rakuten surveys.

We recently sent you an invitation to the Special "3-Year Survey".

**Please participate in our study – this study will be closing soon!**

You are part of a special group of participants across Japan (and in 28 other countries throughout the world) that are invited to participate in this research that will help policymakers and health professionals in forming future policies and programs on tobacco.

The International Tobacco Control (ITC) Japan Survey is a research of people's knowledge, attitudes, and behaviours about tobacco products. Both "users" and "non-users" of tobacco products are needed for this study — a survey that will be conducted once a year over three years.

This is a continuing research study that involves:
- participating in this 35-minute survey during the next few days, and
- if you complete the survey this time, being invited to participate in follow-up surveys every 12-18 months.

If you qualify and fully complete this survey, it's extremely important that you come back and join our Special "3-Year Survey" again in 12-18 months!

We'd appreciate you taking the survey at your earliest convenience by using the web address below.  If you have already completed the survey, thank you and please discard this email.

Please join the below survey.

- **Survey URL:  [LINK TO SURVEY]**
  - Answering a survey can only be done once per person.

- Survey length: About 35 minutes. (Actual length may vary depending on contents)
  - Incentive: [**Rakuten min – max incentive**].

- Closing Date: Once the required number of survey completions is reached.
- Incentive Awarded Date：Max 60 days after completing the survey
  - Depending on the survey, it may take longer than the above timing.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Please take note while answering surveys:

- Please make sure to review each question and answer choice carefully to make sure your answers are accurately selected.

- We suggest the below internet environment: [xx]

- We strictly adhere to the Privacy Policy in the Member Agreement. Please do not disclose any information you acquired in the survey to any third party, including posts to bulletin boards and blogs.

- Surveys that were emailed more than 2 weeks ago will not appear on your My Page. Please click the URL to answer the survey.
    - If the URL is cut off based on your email setting, please copy the whole URL and paste it into your browser address bar.
    - You may not enter the survey if the survey is already closed.

- You will not be able to reply to this email.

- Please go to FAQ for any questions.

**APPENDIX 3. DETAILED SUMMARY OF THE WAVE 1 ITC JAPAN SURVEY COMPLIANCE WITH CHECKLIST FOR REPORTING RESULTS OF INTERNET E-SURVEYS (CHERRIES)**

| ✂ | **Checklist for Reporting Results of Internet E-Surveys (CHERRIES)** |
|---|---|
| *Item Category & Checklist Item* | *ITC Japan Wave 1 (JP1) Survey compliance with CHERRIES* |
| **Design** | |
| Target population | The ITC Japan Wave 1 (JP1) sample was designed to be nationally representative of Japanese cigarette smokers, heated tobacco product (HTP) users, dual users (of cigarettes and HTP products), and non-users. |
| Sampling frame | The ITC JP1 sample comprised the following subgroups: <br><br> <table><tr><td>1) Cigarette (only) smokers</td><td>Smokes cigarettes at least monthly AND (uses HTP not at all or less than weekly).</td></tr><tr><td>2) HTP (only) users</td><td>(Smokes cigarettes not at all or less than monthly) AND uses HTP at least weekly</td></tr><tr><td>3) Cigarette/HTP dual users</td><td>Smokes cigarettes at least monthly AND uses HTP at least weekly</td></tr><tr><td>4) Non-users</td><td>(Smokes cigarettes not at all or less than monthly) AND (uses HTP not at all or less than weekly)</td></tr></table> <br> All of the subsample groups were recruited from Rakuten Insight's Japan panel(s). Rakuten Insight provided the following description of their panel(s):  The JP1 survey was conducted with Rakuten's proprietary online panel in Japan. The online panel is actively managed in-house with a dedicated panel management team in Tokyo, and utilized for market research purposes only. Recruitment for the panel is conducted on a daily basis, tapping into users of Rakuten services (e.g. e-commerce, credit cards, insurance, mobile services, etc.) who collectively comprise of roughly 80% of the Japanese population, as well as other online resources such as affiliates, email and banner recruits in order to maintain a panel as consistent as possible with the general population. Panelists are pre-profiled with a series of questions which in turn can be used as pre-targeting variables (e.g. smoking, HTP usage, etc.). Panelists receive email invitations and also have the option of logging into their proprietary panel site to access the survey they are invited to participate in. Details available at: https://insight.rakuten.com/. |
| Sample | 4684 residents of Japan aged 20y and older were surveyed from the following user groups: <br> • 3306 cigarette smokers [smoked cigarettes at least monthly and (used HTP not at all or less than weekly)], <br> • 207 HTP users [(smoked cigarettes not at all or less than monthly) AND used HTP at least weekly], <br> • 555 cigarette/HTP dual users [smoked cigarettes at least monthly AND used HTP at least weekly], and |

| ![scissors icon] | **Checklist for Reporting Results of Internet E-Surveys (CHERRIES)** |
|---|---|
| *Item Category & Checklist Item* | *ITC Japan Wave 1 (JP1) Survey compliance with CHERRIES* |
| | • 616 non-users [(smoked cigarettes not at all or less than monthly) AND (used HTP not at all or less than weekly)]. |
| **IRB (Institutional Review Board) approval and informed consent process** | |
| IRB approval | All survey procedures and materials were cleared by a University of Waterloo Research Ethics Committee. |
| Informed consent | All participants were part of a Rakuten Insight panel and had an existing agreement with Rakuten Insight to complete surveys for awards. For the ITC JP1 Survey, pre-identified panelists (who may have met survey criteria) were invited by email invitation to complete the survey. The email invitation described the topic of the survey and that the overall project involved one or more follow up survey(s) in the future, the length of the survey, and the incentive value.<br><br>Upon entering the survey, the first few screens reiterated the information above and also explained who was conducting the research, that respondents could skip questions or withdraw, that data were strictly confidential, and provided the ethics contact information. |
| Data protection | No personal identifying information data for panelists was stored in the survey data.  All survey response data files were confidential and maintained on secure servers.  Data were transferred using the University of Waterloo secure system.<br>Rakuten employed strict data protection & security measures to safeguard panelist information. Examples include protecting data by Encryption Communication Technologies (SSL), utilizing Site Authentication Systems to prevent unauthorized access, appointing privacy information officer to manage panelist information, secure data transfer, and adherence to local laws with respect to privacy and confidentiality as stipulated in the Japan Marketing Research Association's guidelines of personal information protection and codes of conduct. |
| **Development and pre-testing** | |
| Development and testing | The survey content and logic was developed by a team of international tobacco control and survey design experts through a structured and iterative process of consultation and revision.  Survey content was developed in English and then translated into Japanese.  The Japanese translation was reviewed and verified by bilingual members of the research team. The survey specifications were then sent to an experienced ITC Survey Management team for operationalization and further extensive systematic review and refinement to ensure survey logic, question wording, response options, and all other survey elements were refined and cross-referenced for consistency, clarity, and accuracy. |

| ✂ | **Checklist for Reporting Results of Internet E-Surveys (CHERRIES)** |
|---|---|
| *Item Category & Checklist Item* | *ITC Japan Wave 1 (JP1) Survey compliance with CHERRIES* |
| | The ITC JP1 Survey was programmed using Confirmit software by Rakuten Insight via a collaborative process between the programmer/fieldwork team and the investigator team to refine survey content as needed to ensure a good survey experience and accurate data collection. The programmed instrument was then systematically and comprehensively tested for usability and technical functionality by the survey firm, as well as by the ITC team. |
| **Recruitment process and description of the sample having access to the questionnaire** | |
| Open survey versus closed survey | The survey was a closed, password-protected survey, in which the respondents could access only their own unique survey record via a predetermined link and/or accessed via the members' private account on the Rakuten Panel site. Each record was associated with a unique ID number. Email invitations included the name of the intended respondent. |
| Contact mode | Panelists were contacted in the standard method (standard Rakuten Insight survey invitation with some minor adaptations for the ITC JP1 study) from Rakuten Insight. |
| Advertising the survey | Only panelists who were pre-identified as being candidates for the survey were invited via a study-specific email invitation. |
| **Survey administration** | |
| Web/E-mail | The ITC JP1 Survey was a web survey hosted by Rakuten Insight on their secure server. |
| Context | The survey could only be accessed by participants who had been specifically invited to participate. All of the persons invited to the survey were Rakuten Insight Japan panelists. |
| Mandatory/voluntary | The survey was voluntary. Up to two reminder emails were sent to those who had not submitted their completed survey in order to maximize the response. |
| Incentives | Panelists were given the Rakuten Insight standard number of points upon completion of the survey, plus bonus points the equivalent to $3 USD. |
| Time/Date | Data collection: February 3- March 2, 2018 |
| Randomization of items or questionnaires | None of the questionnaire items were randomized. This was intentional to prevent any differential priming of respondents. |
| Adaptive questioning | The ITC JP1 Survey used adaptive questioning based on information that the respondent had provided during the survey or to ITC/Rakuten Insight. |
| Number of Items | For most of the survey, one item (i.e., question) was asked per page, with the exception of 'question series'. Each 'question series' consisted of multiple items on one screen (i.e., as grids or a scroll-down format for mobile devices). |

| ✂ | **Checklist for Reporting Results of Internet E-Surveys (CHERRIES)** |
|---|---|
| *Item Category & Checklist Item* | *ITC Japan Wave 1 (JP1) Survey compliance with CHERRIES* |
| Number of screens (pages) | The **TOTAL** number of screens (pages) applicable to the JP1 web survey was over 257 screens.  However, no respondent would ever have been exposed to all of these screens due to routing and filtering of questions for different samples, user types, and response patterns throughout the survey. Each respondent would have seen significantly fewer screens in total, and the number of screens seen across respondents would have varied significantly depending on their sample source, user types, and response patterns. |
| Completeness check | Respondents had to select a response to every survey item in order to progress to the next screen but were able to choose 'Prefer not to answer' for any question.  The survey contained essential questions that were necessary to determine eligibility.  If a respondent was unwilling or unable to answer an essential question, then the respondent was shown a note explaining that they would be unable to continue the survey if they did not provide a response. This explanatory note appeared on the same screen as the essential question. |
| Review step | Respondents were able to review and change their answers at any point, up until they formally submitted their survey.  At the end of the survey, the respondent was prompted to hit the submit button in order to submit their completed responses. |
| **Response rates** | |
| Unique site visitor | A unique visitor was defined based on the respondent's unique ID number. Each record had a unique ID number and was assigned to one predetermined respondent. |
| View rate (Ratio of unique survey visitors/unique site visitors) | Not provided. |
| Response rate | 45.1% |
| Cooperation rate | 96.3% |
| **Preventing multiple entries from the same individual** | |
| Cookies used | Cookies were not used to identify multiple entries, but they were used as part of the survey software. |
| IP check | The IP address of the respondent computer was not used to identify potential duplicate entries from the same user.  Respondents could access their own unique survey record only (via direct link with embedded unique ID number). Respondents were allowed to stop and re-start the survey during the period that fieldwork was open up until the point of formally choosing to submit their data (by choosing this option at the end of the survey). |
| Log file analysis | The log file was not used to identify multiple entries. |

| ✂ | **Checklist for Reporting Results of Internet E-Surveys (CHERRIES)** |
|---|---|
| *Item Category & Checklist Item* | *ITC Japan Wave 1 (JP1) Survey compliance with CHERRIES* |
| Registration | As described in the "IP check" field above, respondents could access their own unique survey record only (via direct link with embedded unique ID number). Respondents were allowed to stop and resume the survey during the period that fieldwork was open up until the point of formally submitting their data (by choosing this option at the end of the survey). Thus, respondents were able to return to where they last left off, unless there was a specific section in the survey where they can over-ride their previous answers. |
| **Analysis** | |
| Handling of incomplete questionnaires | Only completed surveys were included in the final data set released for analyses. |
| Questionnaires submitted with an atypical timestamp | Two criteria were used to assess poor-quality data: 1) Seconds per question (SecperQ) 2) % of responses that were either Refused or Don't Know (%RDK). 1497 respondents were deleted from the final dataset due to very extreme values for both of these variables: SecperQ less than or equal to the 20th percentile of their user-group, and %RDK greater than the 85th percentile of their user-group. Please refer to Section 5.3 in the ITC JP1 Technical Report (https://itcproject.org/technical-report/?country=Japan) for further details on criteria for poor data quality. |
| Statistical correction | Cross-sectional survey weights for different analyses were constructed for the final data set. For the cross-sectional weights, respondents were first divided into four broad user groups: 1) cigarette-only users, 2) HTP-only users, 3) HTP/cigarette dual users, and 4) non-users. The weights were then calibrated on the following cross-tabs using a raking algorithm: user group x gender, user group x age group, user group x geographic region. This weight calibration was done using data from the 2017 Japan Society and Tobacco Internet Study (JASTIS).[1] Please refer to Appendix 4 in the ITC JP1 Technical Report (https://itcproject.org/technical-report/?country=Japan) for further details. |

**References:**

**[1]** Tabuchi T, Shinozaki T, Kunugita N, Nakamura M, Tsuji I. Study Profile: The Japan "Society and New Tobacco" Internet Survey (JASTIS): A Longitudinal Internet Cohort Study of Heat-Not-Burn Tobacco Products, Electronic Cigarettes, and Conventional Tobacco Products in Japan. J Epidemiology. 2018 doi=10.2188/jea.JE20180116.

# Introduction to the Sampling Weights of the International Tobacco Control (ITC) Japan Wave 1 Survey

C. Boudreau[1,2] M. Yan[2,3] and G. Li[2,3]

This short document describes the various cross-sectional sampling weights available for the ITC Japan wave 1 Survey. It also provides some guidance on which set of weights should be used depending on the analysis they are performing. A more detailed document will be created later this year.

# Contents

# 1   Cross-sectional sampling weights

Eight sets of cross-sectional weights were computed at wave 1 of the ITC Japan Survey (summarized in table 1):

1- Variable `aWT23S100v` contains the wave 1 cross-sectional inflation weights for all 4615 respondents interviewed at wave 1. This total excludes the 229 respondents that were deemed to be fraudulent (also referred to as speeders in some documentation).

---

[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 May 22, 2018 (with a minor update to the appendix on Mar 5, 2020)

| Weight | Variable Names |
|---|---|
| Wave 1 cross-sectional inflation weights for all respondents | aWTS23100v |
| Rescaled wave 1 cross-sectional inflation weights for all respondents | aWTS23101v |
| Rescaled wave 1 cross-sectional weights for cigarette smokers | aWTS23201v |
| Rescaled wave 1 cross-sectional weights for HNB users | aWTS23701v |
| Rescaled wave 1 cross-sectional weights for dual users | aWTS23401v |
| Rescaled wave 1 cross-sectional weights for quitters | aWTS23501v |
| Rescaled wave 1 cross-sectional weights for all tobacco users | aWTS23601v |
| Rescaled wave 1 cross-sectional weights for non-users | aWTS23801v |

Table 1: List of the available cross-sectional sampling weights for wave 1 of the ITC Japan Survey

| User group$^\dagger$ | $n$ |
|---|---|
| Cigarette smokers | |
|     Cigarette only | 3288 |
|     Dual users | 549 |
|     Total | 3837 |
| HNB only users | 74 |
| Quitters | |
|     HNB users | 90 |
|     Do not use HNB | 78 |
|     Total | 168 |
| Non-users | 536 |
| Total | 4615 |

$^\dagger$ See cross-sectional inflation weights (variable aWTS23100v)
for the description of those user groups

Table 2: Respondents by user group.

These weights were computed by dividing the respondents into 6 user groups (variable aUserGrp[1] in the dataset): *i*) cigarette only users, *ii*) dual users, *iii*) HNB only users, *iv*) short term quitter (i.e., quit cigarette smoking within the last two years) using HNB at least weekly, *v*) short term quitter using HNB less than weekly or not at all, and *vi*) non-users & long-term quitters; see table 2 for the number of respondents in each of the 6 user groups. All respondents had to be 20 years or older at the time they completed the survey. To be categorized as a cigarette smoker (group *i*), the respondent had to smoke

---

[1] Where 1 = cigarette only, 2 = HNB only, 3 = dual, 4 = non-user, 5 = recent quitter using HNB at least weekly and 6 = recent quitter using HNB less than weekly; see appendix.

cigarettes at least monthly and at least 100 cigarettes in their lifetime, and use HNB less than weekly or not at all. To be categorized as a dual user (group *ii*), the respondent had to smoke cigarettes at least monthly (and at least 100 cigarettes in their lifetime) and use HNB at least weekly at the time of data collection. To be classified as heat-not-burn (HNB) only user (group *iii*), the respondent had to use HNB at least weekly and smoke cigarettes less than monthly or not at all. To be classified as a short term quitter (groups *iv* and *v*), the respondent had to had quit cigarette smoking within the last two years, and use HNB at least weekly (group *iv*), or use HNB less than weekly or not at all (group *v*). Finally, to be classified as a non-user/long-term quitter (group *iv*), the respondent had to smoke cigarettes less than monthly or not at all, use HNB less than weekly or not at all, and either never have been a cigarette smoker (or smoked les than 100 cigarettes in their lifetime) or quit more than 2 years ago. See appendix for computer code use to create the 6 user groups.

In addition to the 6 user groups, respondents were further subdivided based on gender, age group (i.e., [20, 25), [25, 35), [35, 45) , [45, 55) & [55, 80]), and education (low, medium & high). This yielded the following 3 cross-tabs: user group × gender, user group × age group, and user group × education. Note that some cells in the user group × age group and user group × education cross-tabs were collapsed because they contained too few respondents. In particular, all age groups were collapsed into a single group for HNB only users (group *iii*); likewise for education. Finally, respondents were divided based on the 8 geographic regions of japan. However, region was not crossed with user groups.

Calibration/target figures (e.g., estimated number of individuals that are dual users) were obtained for each of the 3 cross-tabs and the 8 geographic regions. These calibration figures were obtained using data from the 2017 Japan Society and Tobacco Internet Study (JASTIS). A raking procedure was then applied to calibrate the weights using the above mentioned cross-tabs and geographic regions.

These weights are designed to make respondents in each of the 6 groups representative (with respect to gender, age, education and region) of the corresponding population at the time of wave 1 data collection. For example, the `aWTS23100v` weights of the 549 dual users are designed to make them representative of the Japanese population of dual users at the time of data collection; likewise for the other 5 groups.

If interest lies in a target population that consists of two or more of the 6 user groups, the `aWTS23100v` weights are still appropriate. For example, when studying cigarette smokers, one can simply combine the `aWTS23100v` weights of the 3288 cigarette only users with those of the 549 dual users (for a total of 3288 respondents in the analysis), and assigned a weight of 0 to respondents in the other 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 `aWTS23101v` to `aWTS23801v`) described below were created especially for such multi-country analyses; see section 2.1 for more information on inflation versus rescaled weights.

2- Variable `aWTS23101v` contains the rescaled wave 1 cross-sectional weights for all 4615 respondents. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of those respondents rescaled to sum to sample size (i.e., $n = 4615$). These weights are designed to make these 4615 respondents representative of the adult (20 years and older) Japanese population (cigarette smokers, HNB users, recent quitters, and non-users & long-term quitters) at the time of wave 1 data collection.

It should be noted that tobacco users, quitters and non-users are ultimately distinct populations. Hence, great care must be taken when deciding to analyse them together using the `aWTS23101v` weights. This is probably fine when the goal is to carry out descriptive inference about the joint population of tobacco users, quitters and non-users. However, carrying out analytical inference (e.g., linear regression and logistic regression) from that same joint population is probably much more questionable.

3- Variable `aWTS23201v` contains the rescaled wave 1 cross-sectional weights for the 3837 (see table 2) respondents who were cigarette smokers at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of those respondents rescaled to sum to sample size (i.e., $n = 3837$). These weights are designed to make these 3837 cigarette smokers representative of the Japanese population of cigarette smokers at the time of wave 1 data collection.

4- Variable `aWTS23701v` contains the rescaled wave 1 cross-sectional weights for the 713 (see table 2) respondents who were HNB users at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of those respondents rescaled to sum to sample size (i.e., $n = 713$). These weights are designed to make these 713 HNB users representative of the Japanese population of HNB users at the time of wave 1 data collection.

5- Variable `aWTS23401v` contains the rescaled wave 1 cross-sectional weights for the 549 (see table 2) respondents who were dual users at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of those respondents rescaled to sum to sample size (i.e., $n = 549$). These weights are designed to make these 549 dual users representative of the Japanese population of dual users at the time of wave 1 data collection.

6- Variable `aWTS23501v` contains the rescaled wave 1 cross-sectional weights for the 168 (see table 2) respondents who were recent quitters (i.e., within the last 2 years) at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of those respondents rescaled to sum to sample size (i.e., $n = 168$). These weights are designed to make these 168 recent quitters representative of the Japanese population of recent quitters at the time of wave 1 data collection.

7- Variable `aWTS23601v` contains the rescaled wave 1 cross-sectional weights for the 4001 (see table 2) respondents who were tobacco users at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of

those respondents rescaled to sum to sample size (i.e., $n = 4001$). These weights are designed to make these 4001 tobacco users representative of the Japanese population of tobacco users at the time of wave 1 data collection.

8- Variable `aWTS23801v` contains the rescaled wave 1 cross-sectional weights for the 536 (see table 2) respondents who were not using any tobacco products or were long-term quitters (i.e., quit more than 2 years ago and not using HNB) at the time of wave 1 data collection. These are simply the wave 1 cross-sectional inflation weights (variable `aWTS23100v`) of those respondents rescaled to sum to sample size (i.e., $n = 536$). These weights are designed to make these 536 respondents representative of the Japanese population of non-users and long-term quitters.

# 2 Additional remarks

## 2.1 Inflation versus rescaled weights

The main reason for rescaling the weights is to facilitate joint analyses involving data from multiple ITC countries. Using data from the 2017 JASTIS used to calibrate the wave 1 cross-sectional inflation weights (i.e., variables `aWTS100v`), there were about 18.3 million cigarette smokers (ages 20 and older) in Japan at the time of wave 1 data collection. Data from the 2016 National Health Interview Survey (NHIS) was used to calibrate the weights of the US sample for wave 1 of the ITC Four Country Tobacco and E-Cigarette (4CE) Survey, which was conducted in 2016. According to the 2016 NHIS, there were about 39.8 million cigarette smokers (ages 18 and older) in the United States at the time of data collection; Hence, any joint analysis using data from ITC Japan and the US sample of the 4CE Survey will be dominated by the US if the inflation weights (i.e., variable aWTS23100v) are used.

On the other hand, the various rescaled weights sum to the appropriate sample size, as described above; and likewise for the 4CE Survey. Hence, if the rescaled weights are used, Japan will have a slightly greater impact on the results (the ITC Japan sample of cigarette smokers consists of 3837 respondents; whereas the 4CE sample of US cigarette smokers consists of 2327 respondents), 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 ITC Japan data to other ITC countries.

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 e-cigarette users).

## 2.2 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 ITC Japan data. Hence, we highly recommend that any statistical model includes the following covariates:

- gender (labelled `sex` in the dataset)
- age (labelled `aAge`, continuous, and `aAgeGrp`, categorical, in the dataset)
- user group (labelled `aUserGrp` in the dataset)

Though somewhat less essential, users should also strongly consider adding education (labelled `aDE23312v` in the dataset) to their statistical model(s). The geographic region (labelled `aStrata` in the dataset) should also be used as the stratification variable in the statistical software.

# Appendix: pseudo code

Pseudo code detailing how variable `aUserGrp` (abbreviated `UserGrp`) was created.

```
if {smokes cigarettes at least monthly (ie, kFR309v in (1,2,3))} then
   do;
      if {uses HNB at least weekly (ie, HN309v in (1,2))} then
         UserGrp=3; /* dual */
      else then UserGrp=1; /* cig only */
   end;

else if {non-smoker or quitter (ie, FR309v=9)} and {QA23439=n/a} then
   do;
      if {uses HNB at least weekly (ie, HN309v in (1,2))} then
         UserGrp=2; /* HNB only */
      else then UserGrp=4; /* non-user */
   end;

else if {non-smoker or quitter (ie, FR309v=9)}
         and {quit within the last 2 years (ie, QA439 in (1:4))} then
   do;
      if {uses HNB at least weekly (ie, HN309v in (1,2))} then
         UserGrp=5; /* quitter using HNB at least weekly */
      else then UserGrp=6; /* quitter using HNB less than weekly */
   end;

else if {non-smoker or quitter (ie, FR309v=9)}
         and {quit more than 2 years ago (ie, QA439 in (5,6,7,88,99))} then
   do;
      if {uses HNB at least weekly (ie, HN309v in (1,2))} then
         UserGrp=2; /* HNB only */
      else then UserGrp=4; /* non-user */
   end;
```

Note: the response option for variable QA439 changed between waves 1 and 2; at wave 1, response options 1–4 corresponds to those who quit within the last 2 years.

**APPENDIX 4.2: BENCHMARKS FOR ITC JAPAN WAVE 1 SURVEY WEIGHTS**

| Survey used for weights benchmarks | Comments and limitations |
|---|---|
| The Japan Society and New Tobacco Internet Survey (JASTIS) 2017 internet survey of a nationally representative panel of Japanese residents: sample size 5897 (aged 15-69) | Internet survey may not be as representative as address-based frames.  JASTIS included respondents aged as old as 69 years, whereas the JP1 Survey included older respondents.   The JASTIS Study Profile [1] lists other strengths and limitations of the study, which include small sample sizes for HTP only, recent quitters (both using and not using HTP). |

**References:**
[1] Tabuchi T, Shinozaki T, Kunugita N, Nakamura M, Tsuji I. Study Profile: The Japan "Society and New Tobacco" Internet Survey (JASTIS): A Longitudinal Internet Cohort Study of Heat-Not-Burn Tobacco Products, Electronic Cigarettes, and Conventional Tobacco Products in Japan.
J Epidemiology. 2018 doi=10.2188/jea.JE20180116.

**APPENDIX 5: METHODS FOR IMPUTING CIGARETTES PER DAY**

PROCEDURES FOR JAPAN WAVE 1.5 CPD DATA IMPUTATION

**Introduction**

The ITC Japan Project was conducted from February 3 to March 2, 2018. The ITC Japan Survey was designed to provide data on smokers' knowledge, attitudes, beliefs, perceptions, behaviors, and use patterns associated with HTP products from the most important market for these products. Unfortunately, the **cigarettes consumed per day** (**CPD,** collected via variables FR216, FR226, and FR236) for smokers and ET634 (with a revised translation from the JP1 version) were not included in the Wave 1 Survey. Thus, the research team conducted the ITC Japan 1.5 (JP1.5) Survey to collect the missing data.

- The JP1.5 sample was a subset of respondents who completed the JP1 Survey. Panelists in the sample were subsequently re-surveyed and asked 1-2 additional questions to be answered as though it were the time of the JP1 Survey.
- The JP1.5 sample is summarized in Table A5.1, below.
- The inclusion criteria were: adults aged 20 y or older, current Rakuten panelists who are eligible to answer 1 or 2 of FR216, FR226, FR236, or ET634 (as determined by previous responses to the JP1 Survey).
- Exclusion criteria were: absence of any inclusion criterion.

**Table A5.1. The ITC Japan Wave 1.5 Sample as a function of 'cigarettes per day' and ET634.**

| Questions respondents will be asked | Number of JP1 respondents eligible for question(s) | Number of JP1 respondents answered the question(s) |
|---|---|---|
| CPD and ET634 | 2347 (FR216 =2193, FR226 =135, FR236 =19) | 1885 (including 1 refused and 9 don't know) |
| Only CPD | 1514 (FR216 =1456, FR226 =50, FR236 =8) | 1258 (including 1 refused and 14 don't know) |
| Only ET634 | 429 | 345 (including 1 refused and 9 don't know) |
| Total | 4,290 | 3488 |

**Response Rates**
- Of the eligible sample, 81.3% completed the Wave 1.5 Japan Survey.
- 3143 out of 3861 (81.4%) respondents answered the CPD question, with 25 (0.65%) refused and don't know.

**Imputing missing values for cigarettes consumed per day (CPD)**
- CPD is a key measure of smoking prevalence in the ITC survey. Thus, missing data/non-response at Wave 1.5 for cigarettes per day were imputed. Here are the step by step details.

**Procedures**
- CPD is derived by FR216 (Daily smokers), FR226 (Weekly smokers) and FR236 (Monthly smokers) from JP1.5 data. Respondents who answered 8888 (Refused) and 9999 (Don't know) to FR216, FR226 or FR236 are set as missing. And we name the derived CPD as FR245v.

- For the ITC Japan Wave 1 data, the missing data of CPD is due to the follow-up attritions and the JP1.5 non-responses. Thus, the research team first examined the variables associated with attrition.

**Attrition models**
- There were 3861 respondents eligible to answer the CPD question while there were 3143 who participated the follow-up survey. Thus, the research team defined an indicator variable called inM15 that inM15=1 if the respondent answered the question and inM15=0 if he/she dropped out. Then PROC UNIVARIATE procedures with inM15 as the response variable and each of the survey questions as the explanatory variable were used to filter out all the possible survey questions associated with the attrition.
- The following explanatory variables were then added to PROC GLMSELECT procedure to run a stepwise selection. Some of the following survey questions were not asked to all respondents, so an extra response option of "non-applicable", or "non-users", etc., was added to the data to replace the missing values of each variable.

```
Sex     /* respondent's gender */
AGEGRP /* respondent's age group*/
rSmoke /* smoking status at recruitment*/
JPQregionv /* respondent's region (derived) */
HN23106 /* ever used HNB(if heard)*/
HN23307 /*consider self-addicted to HNB*/
HN23247 /*HNB brand chosen bec same as friends*/
HN23235 /*your HNB variety/flavour how harmful vs others*/
HN23810 /*bought HNB device kit with supplies*/
HN23762 /*freq seen HNB use indoors where cigs banned*/
HN23774 /*you use HNB: at bar, pub*/
HN23411 /*noticed HNB ad: email or texts*/
HN23415 /*noticed HNB ad: bars or pubs*/
FR23353 /*RYO reduces amount I smoke*/
BR23358 /*freq, crush the flavour pellet*/
LM23331 /*lights smoother on throat and chest*/
CH23960 /*used insurance-covered cessation service LQA*/
SM23949 /*used any other SSM at last QA*/
SM23930 /*used other NRT at last QA*/
KN23265 /*SHS causes harm to kids, elderly*/
WL23221 /*freq, cig WL stopped you from smoking*/
WL23313 /*avoid cig WL in any way*/
WL23425 /*cig WL alarm or calm you*/
ET23624 /*You smoked at workplace*/
ET23882 /*support smoking ban in restaurants*/
AD23214 /*noticed tob ads in email or texts*/
AD23217 /*noticed tob ads at bars or pubs*/
PU23660 /*to save money on cig: other*/
AD23701 /*noticed anti-smoking info at all*/
PS23328 /*society's attitude towards smoking cigs*/
IN23220 /*tobco should use plain packages for cigs*/
IN23337 /*support ban on menthol*/
PU23680 /*gov't should increase tax on cig*/
PR23220 /*smoking-related disease prob: keep smoking*/
PR23313 /*worried smoking will damage your health*/
RE23255 /*SH vapour harm vs SH cig smoke*/
HN23480 /*# HNB users of five closest friends*/
DI23712 /*how often drink alcohol*/
DI23703 /*number of alcohol drinks on typical day*/
DE23311 /*highest level of education*/
DI23248 /*partner/ spouse thinks you should quit cigs*/
HN23490 /*does spouse/ partner use HNB*/
ME23449 /*# online surveys completed recently*/
TTHNB   /* Time to the first HNB */
SB23012v /* Time to the first cig (derived)*/
```

```
            ▪   Device  /* Device using */
```

- The outputs of PROC GLMSELECT procedure suggests a poor fit of attrition models, so the research team decided to group the Wave 1 smoker data into 3 users groups. They were *cigarette users* who currently used HNB less than monthly or who tried before, *dual users* who currently used cigarettes and HTP, and the *pure cigarette users* who never tried or heard of HTP before.  These 3 mutually exclusive groups provided a better fit of the GLMSELECT models.

**Table A5.2.1 Variables associated with attrition between JP1 and JP1.5.**

| | # | Groups | | |
|---|---|---|---|---|
| | | *pure cigarette users*<br>N = 2342 | *cigarette users (HTP triers)*<br>N = 964 | *dual users*<br>N = 555 |
| **Variables associated with attrition** | 1 | AGEGRP | AGEGRP | AGEGRP |
| | 2 | JPQREGIONV | HN23307 | HN23307 |
| | 3 | HN23762 | HN23810 | BR23358 |
| | 4 | SM23949 | FR23353 | SM23949 |
| | 5 | KN23265 | SM23949 | AD23217 |
| | 6 | WL23221 | AD23214 | DI23712 |
| | 7 | WL23425 | PS23328 | ME23449 |
| | 8 | PU23660 | DI23712 | Device |
| | 9 | PS23328 | ME23449 | |
| | 10 | IN23220 | Device | |
| | 11 | IN23337 | | |
| | 12 | HN23480 | | |
| | 13 | Device | | |

* All these variables were added to their corresponding CPD imputation models regardless whether it significantly associates with CPD or not.

**CPD Imputation Models**
- There were 3861 respondents are eligible to answer the CPD question while there were only 3143 respondents have valid responses. So the research team aimed to impute values for the 718 missing.
- Similar to the attrition models, the team first use PROC UNIVARIATE procedures with CPD (FR245v) as the response variable and each of the survey questions as the explanatory variable.  Then determined all the possible survey questions associated with the CPD numbers. Some borderline explanatory variables were also included in the lists (i.e., with p-value slightly larger than 0.05).  To be consistent, the team used the same 3 mutually exclusive groups: *cigarette users* who tried HTP, the *dual users*, and the *pure cigarette users.*
- The following explanatory variables were then added to PROC GLMSELECT procedure to run a stepwise selection. As defined previously, an extra response option of "non-applicable", or "non-users", etc., was added to the data to replace the missing values of each variable.

```
            ▪   sex     /*Respondent's gender*/
```

- AGEGRP   /*age at recruitment (categorical)*/
- FR23309V /*smoking status (derived)*/
- HN23309V /*heat-not-burn status (derived)*/
- HN23012v /*Time to the first HNB*/
- HN23506 /*why use HNB: less harmful than cigs*/
- HN23509 /*why use HNB: for the taste*/
- HN23510 /*why use HNB: use in smokefree areas*/
- HN23532 /*why use HNB: Drs, scientists use them*/
- HN23514 /*why use HNB: look cool*/
- HN23528 /*why use HNB: attractive packaging*/
- HN23516 /*why use HNB: less harmful to others */
- HN23524 /*why use HNB: help manage stress*/
- HN23519 /*why use HNB: enjoyment*/
- HN23566 /*how satisfying is HNB vs cig*/
- HN23243 /*HNB brand chosen for taste*/
- HN23244 /*HNB brand chosen for design*/
- HN23249 /*HNB brand chosen bec advertising*/
- HN23248 /*HNB brand chosen bec media people use it*/
- HN23235 /*your HNB variety/flavour how harmful vs others*/
- HN23621 /*noticed HNB WL*/
- HN23622 /*read or look closely at HNB WL*/
- HN23810 /*bought HNB device kit with supplies*/
- HN23762 /*freq seen HNB use indoors where cigs banned*/
- HN23731 /*rules for using HNB inside your home*/
- HN23773 /*you used HNB: at restaurant, café*/
- HN23771 /*you use HNB: outdoors at home*/
- HN23776 /*you use HNB: in other people's homes*/
- HN23777 /*you use HNB: in MUH public areas*/
- HN23404 /*noticed HNB ad: on radio*/
- HN23406 /*noticed HNB ad: on posters or billboards*/
- HN23410 /*noticed HNB ad: regular mail*/
- HN23412 /*noticed HNB ad: on websites or social media*/
- HN23415 /*noticed HNB ad: bars or pubs*/
- HN23419 /*noticed HNB ad: at sporting events*/
- HN23421 /*noticed HNB logo on clothing, etc.*/
- HN23423 /*noticed HNB free samples*/
- HN23702 /*support complete ban on HNB POP displays*/
- HN23596 /*how worried: HNB will damage your health*/
- FR23333 /*mainly factory-made vs RYO*/
- FR23353 /*RYO reduces amount I smoke*/
- PS23243 /*harsh smoke is more dangerous*/
- LM23705 /*menthol smoother on throat and chest*/
- LM23703 /*menthol are less harmful*/
- SB23012v /*Time to the first cigs*/
- SB23031 /*how addicted to cigs*/
- QA23561 /*# quit attempts since {time}*/
- QA23671 /*last quit, stop suddenly or cut down*/
- SM23949 /*used any other SSM at last QA*/
- KN23221 /*smoking causes stroke*/
- KN23265 /*SHS causes harm to kids, elderly*/
- KN23264 /*smoking causes addiction, damage to underage SM*/
- KN23251 /*SHS causes lung cancer in NS*/
- WL23201 /*freq, notice cig WL*/
- WL23211 /*freq, read or look closely at cig WL*/
- WL23221 /*freq, cig WL stopped you from smoking*/
- WL23421 /*cig WL make you more likely to quit*/
- WL23427 /*cig WL make you worried*/
- WL23545 /*support policy of graphic cig WL*/
- ET23221 /*describe smoking at home*/
- ET23421 /*rules in local bars*/
- ET23436 /*smoked indoors in bar*/
- ET23536 /*you smoked indoors in restaurant*/
- ET23621 /*smoking policy where you work*/
- ET23634 /*people smoking at workplace*/
- ET23625 /*you smoked indoors at work*/
- ET23888 /*support smoking ban in bars and pubs*/
- ET23819 /*support smoking ban in workplaces*/

```
▪  ET23792 /*support smoke-free Tokyo Olympics*/
▪  AD23202 /*noticed tob ads on websites or social media*/
▪  AD23217 /*noticed tob ads at bars or pubs*/
▪  AD23629 /*support complete ban on POP displays*/
▪  AD23627 /*support complete ban on ads in shops*/
▪  SO23225 /*last purchase of cigs where*/
▪  PU23201 /*last purchase in what form*/
▪  PU23653 /*to save money on cig: consider quitting*/
▪  PU23657 /*to save money on cig: reduce # smoked*/
▪  PU23656 /*to save money on cig: buy bulk*/
▪  AD23736 /*anti-smoking info in shops*/
▪  BQ23111 /*sure would succeed at quitting cigs*/
▪  BQ23141 /*plan to quit smoking*/
▪  BQ23201 /*why quit: concern for personal health*/
▪  BQ23203 /*why quit: effect on non-smoker health*/
▪  BQ23211 /*why quit: restrictions at work*/
▪  BQ23214 /*why quit: friends and family disapprove*/
▪  BQ23239 /*why quit: availability of HTP products*/
▪  PS23215 /*do again, not start smoking*/
▪  PS23223 /*smoking is important part of life*/
▪  PS23328 /*society's attitude towards smoking cigs*/
▪  IN23220 /*tobco should use plain packages for cigs*/
▪  IN23336 /*support for restricting purchase places*/
▪  IN23327 /*support cig ban if other tobacco available*/
▪  PU23680 /*gov't should increase tax on cig*/
▪  PR23101 /*describe your health*/
▪  BQ23301 /*how much benefit if you quit*/
▪  NC23302V /*ever used e-cigs (derived)*/
▪  DI23301 /*overall opinion of smoking cigs*/
▪  HN23560 /*overall opinion of using HNB*/
▪  DI23712 /*how often drink alcohol*/
▪  DI23703 /*number of alcohol drinks on typical day*/
▪  DI23443 /*felt sad in the last week*/
▪  BI23322 /*number 18+ in household*/
▪  BI23324 /*number 18+ who currently smoke*/
▪  DE23211 /*household income*/
▪  DE23311 /*highest level of education*/
▪  DI23245 /*does spouse/ partner smoke*/
▪  DI23248 /*partner/ spouse thinks you should quit cigs*/
▪  ME23450 /*# online smoking surveys completed recently*/
```

- In addition, variables associated with attrition and some possible interaction terms such as age groups by gender, age groups by heaviness of HTP use/heaviness of smoking, gender by smoking status were all added into the models for selection.
- Three selection options of PROC GLMSELECT procedures were used to compare which results make more sense. We use the plots and CV PRESS scores to do the comparisons. Here we call them method A, method B and Method C. (For detailed model comparisons, please see attached Appendix)

**Table A5.2.2 Variables associated with CPD in JP1.5.**

| | # | *pure cigarette users*<br>N = 1963<br>N_used = 1951 | *cigarette users (HTP triers)*<br>N = 749<br>N_used = 740 | *dual users*<br>N = 431<br>N_used = 427 |
|---|---|---|---|---|
| | | | **Groups** | |
| **Variables associated with CPD** | 1 | sex | sex | AGEGRP |
| | 2 | FR23309V | AGEGRP | FR23309V |
| | 3 | HN23731 | FR23309V | HN23012v |
| | 4 | SB23012v | HN23012v | HN23516 |
| | 5 | SB23031 | HN23249 | HN23235 |
| | 6 | KN23251 | HN23421 | SB23012v |
| | 7 | ET23536 | HN23702 | SB23031 |
| | 8 | ET23621 | SB23012v | AD23629 |
| | 9 | PU23201 | WL23421 | BQ23111 |
| | 10 | BQ23141 | ET23634 | Device |
| | 11 | IN23336 | ET23888 | |
| | 12 | DE23211 | ET23819 | |
| | 13 | AGEGRP | BQ23214 | |
| | 14 | HN23480 | BQ23239 | |
| | 15 | | IN23327 | |
| | 16 | | ME23450 | |

\* The order of the above variables are based on the stepwise selection order.

### 5.2.3 Imputation

- The above variables were added to PROC MI procedures which uses methods that incorporate appropriate variability across the multiple times of imputations for the missing values of CPD by groups. Instead of filling in a single value for each missing CPD value, this multiple imputation strategy replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. Here, we have defined that the missing values of CPD are filled in 5 times to generate 5 complete data sets. Only 1 set of the imputed CPD is randomly chosen and then added back to the JP Wave 1 dataset. Investigators can directly use the FR23245v (label = "imputed cigs per day") for their analyses.
- For more statistically analyses where CPD is an important factor, the PROC MIANALYZE procedure is recommended since this procedure combines all the results of the analyses of imputations from the 5 complete data sets and generates valid statistical inferences. Please refer to the SAS manual.
- The imputation results for 3 user groups are listed below:

**Table A5.3. CPD imputation results (based on 5 sets of imputation)**

| FR23245v (CPD with imputation) | Parameters | Groups | | |
|---|---|---|---|---|
| | | *pure cigarette users* N = 2342 | *cigarette users (HTP triers)* N = 964 | *dual users* N = 555 |
| Variance | Between | 0.000451 | 0.008446 | 0.034873 |
| | Within | 0.030114 | 0.088936 | 0.121809 |
| | Total | 0.030654 | 0.099072 | 0.163656 |
| DF | | 1949.3 | 264.84 | 53.249 |
| Relative Increase in Variance | | 0.017956 | 0.113966 | 0.343546 |
| Fraction Missing Information | | 0.017792 | 0.106968 | 0.278896 |
| Relative Efficiency | | 0.996454 | 0.979054 | 0.947168 |
| Mean | | 14.57 | 15.36 | 14.83 |
| Std. Error | | 0.175 | 0.315 | 0.405 |
| 95% Confidence Limits | | (14.23, 14.91) | (14.74, 15.98) | (14.02, 15.65) |
| Minimum | | 14.55 | 15.24 | 14.59 |
| Maximum | | 14.59 | 15.46 | 15.11 |

- The last step, we have derived a categorical variable called FR23250v (label = "Cigs per day (categorical)") based on the FR23245v. It is defined as:

```
0 = '1-10 cigarettes'
1 = '11-20 cigarettes'
2 = '21-30 cigarettes'
3 = 'more than 31 cigarettes'
```

| Groups | Models | Selected Common variables | Extra Variables by each Model | Variable Labels |
|---|---|---|---|---|
| **Pure Cig Smokers**<br><br>**N = 1963**<br>**N_used = 1951** | | HN23731<br>SB23031<br>KN23251<br>ET23536<br>ET23621<br>PU23201<br>BQ23141<br>IN23336<br>DE23211 | ..................................<br>..................................<br>..................................<br>..................................<br>..................................<br>..................................<br>..................................<br>..................................<br>.................................. | rules for using HTP inside your home<br>how addicted to cigs<br>SHS causes lung cancer in NS<br>you smoked indoors in restaurant<br>smoking policy where you work<br>last purchase in what form<br>plan to quit smoking<br>support for restricting purchase places<br>household income |
| | A | | ET23221<br>ET23436<br>ET23625<br>PS23215<br>sex*AGEGRP<br>FR23309v*SB23012v | describe smoking at home<br>you smoked indoors in bar<br>you smoked indoors at work<br>do again, not start smoking<br>sex and ageGrp interaction<br>smoking status and time to the first cig interaction |
| | B | | Sex<br>AGEGRP<br>FR23309v<br>SB23012v<br>ET23625<br>PS23215<br>IN23220<br>IN23327<br>FR23309v*SB23012v | Sex<br>Age group<br>Smoking status<br>time to the first cig<br>you smoked indoors at work<br>do again, not start smoking<br>tobco should use plain packages for cigs<br>support cig ban if other tobacco available<br>smoking status and time to the first cig interaction |
| | C | | sex<br>AGEGRP<br>FR23309v<br>SB23012v<br>HN23480 | Sex<br>Age group<br>Smoking status<br>time to the first cig<br># HTP users of five closest friends |

50

## Appendix 5.1

### PURE CIG SMOKERS -- PLOTS FOR MODEL C:

| Groups | Models | Selected Common variables | Extra Variables by each Model | Variable Labels |
|---|---|---|---|---|
| **Cig Smokers tried HTP Previously or Occasionally**<br><br>**N = 749**<br>**N_used = 740** | | sex<br>HN23249<br>HN23421<br>ET23888<br>BQ23214<br>BQ23239 | …………………………..<br>…………………………..<br>…………………………..<br>…………………………..<br>…………………………..<br>………………………….. | Gender<br>HTP brand chosen bec advertising<br>noticed HTP logo on clothing, etc<br>support smoking ban in bars /pubs<br>why quit: friend/family disapprove<br>why quit: HTP prod availability |
| | A | | KN23221<br>WL23421<br>ET23536<br>ET23621<br>PU23201<br>PS23328<br>PR23101<br>ME23450<br>HN23012v*AGEGRP<br>FR23309V*SB23012v | smoking causes stroke<br>WL make you more likely to quit<br>you smoked indoors in restaurant<br>smoking policy where you work<br>last purchase in what form<br>society's attitude towards smoking<br>describe your health<br>#online smoking surveys completed<br>Time to the first HTP and age group interaction<br>Smoking status and time to the first cig |
| | B | | AGEGRP<br>FR23309V<br>HN23012v<br>SB23012v<br>ET23536<br>ET23621<br>PU23201<br>PS23328 | Age group<br>Smoking status<br>Time to the first HTP product<br>Time to the first cig<br>you smoked indoors in restaurant<br>smoking policy where you work<br>last purchase in what form<br>society's attitude towards smoking |
| | C | | AGEGRP<br>FR23309V<br>HN23012v<br>SB23012v<br>HN23702<br>WL23421<br>ET23634<br>ET23819<br>IN23327<br>ME23450 | Age group<br>Smoking status<br>Time to the first HTP product<br>Time to the first cig<br>support complete ban on HTP POP displays<br>WL make you more likely to quit<br>people smoking at workplace<br>support smoking ban in workplaces<br>support cig ban if other tobacco available<br>#online smoking surveys completed |

Fit Criteria for FR23245v



Progression of Average Squared Errors by Role for FR23245v

| Groups | Models | Selected Common variables | Extra Variables by each Model | Variable Labels |
|---|---|---|---|---|
| **Dual Users N = 431 N_used = 427** | | SB23031 AD23629 Device | ……………………… ……………………… ……………………… | how addicted to cigs support complete ban on POP displays survey device used |
| | A | | HN23012v HN23528 HN23516 WL23545 ET23221 BQ23111 PS23215 BI23322 sex*AGEGRP FR23309V*SB23012v | Time to the first HTP after waking why use HTP: attractive packaging why use HTP: less harmful to others support policy of graphic cig WL describe smoking at home sure would succeed at quitting cigs do again, not start smoking number 18+ in household sex and age group interaction smoking status and time to first cig interaction |
| | B | | sex AGEGRP FR23309V HN23528 HN23248 SB23012v ET23221 PU23657 PU23656 BQ23141 PS23215 BI23322 HN23307 | Gender Age group Smoking status why use HTP: attractive packaging HTP brand chosen bec media people use it Time to the first cig describe smoking at home to save money on cig: reduce # smoked to save money on cig: buy bulk plan to quit smoking do again, not start smoking number 18+ in household consider self-addicted to HTP |
| | C | | AGEGRP FR23309V HN23012v HN23516 HN23235 SB23012v BQ23111 | Age group Smoking status Time to the first HTP after waking why use HTP: less harmful to others your HTP variety how harmful vs others Time to the first cig Sure would success quitting cigs |

Fit Criteria for FR23245v



Progression of Average Squared Errors by Role for FR23245v

**Model Definitions:**
- Model A: `selection = stepwise(select=sl slentry=0.05 slstay=0.05);`
- Model B: `selection = stepwise(select=sl slentry=0.05 slstay=0.05`
  `choose = cv)         stats     = press`
  `            cvMethod  = random(5)`
  `            cvDetails = all`
  `            hierarchy = single stb`
  `            details = all;`
- Model C: `partition fraction(validate=0.2);`
  `            selection = stepwise(select=sl slentry=0.05`
  `            slstay=0.05 choose = cv)`

```
stats    = press
cvMethod = random(5)
cvDetails = all
hierarchy = single stb
details = all;
```

## Appendix 5.2: SAS Codes

Please note that the following SAS codes only include essential information for each of the steps. For the complete codes, please contact DMC.

```sas
/* Attrition models */
data X15a;
      set x15;
      if HN23309v in (4:8) then HN23307=4;
      if HN23309v in (7,8) then HN23247=3;
      if HN23309v in (4:8) then HN23235=4;
      if HN23309v in (1:3) and HN23810 = . then HN23810=3;
      else if HN23309v in (4:8) then HN23810=4;
      if HN23309v = 8 then HN23762 = 5;
      if HN23309v in (4:8) then HN23774=3;
      if HN23309v = 8 then HN23411 = 4;
      if HN23309v = 8 then HN23415 = 4;
      if FR23309v = 9 then FR23353 = 3;
              else if FR23326 in (1,8,9) then FR23353 = 4;
      if FR23309v = 1 then do;
              if SB23012 in (1:4) then SB23012v = 4 - SB23012;
              else if SB23012 in (8,9) then SB23012v = SB23012;
      end;
      else if FR23309v in (2,3) then do;
              if SB23027 in (1:4) then SB23012v = 4 - SB23027;
              else if SB23027 in (8,9) then SB23012v = SB23027;
      end;
      if FR23309v = 9 then BR23358 = 6; /*non-smoker */
              else if BR23358 = . then BR23358 = 7;
      if FR23309v in (4:9) then LM23331=6;
      if FR23309v =9 then CH23960=3;
              else if QA23101 in (2,8,9) then CH23960=4;
              else if QA23561 in (0,8888,9999) then CH23960=5;
      if FR23309v =9 then SM23942=3;
              else if QA23101 in (2,8,9) then SM23942=4;
              else if QA23561 in (0,8888,9999) then SM23942=5;
      if FR23309v =9 then SM23949=3;
              else if QA23101 in (2,8,9) then SM23949=4;
              else if QA23561 in (0,8888,9999) then SM23949=5;
      if FR23309v =9 then SM23930=3;
              else if QA23101 in (2,8,9) then SM23930=4;
              else if QA23561 in (0,8888,9999) then SM23930=5;
              else if SM23920 in (2,8,9) then SM23930=2;
      if FR23309v =9 then WL23221=5;
      if FR23309v =9 then WL23313=3;
      if ET23603 in (2,8,9) then ET23634 = 3;
              else if ET23603 = . then ET23634 = 4;
      if FR23309v =9 then PU23660=3;
      if FR23309v =9 then PR23220=6;
      if FR23309v =9 then PR23313=5;
      if RE23255 = . then RE23255 = 4;
      if HN23309v = 8 then HN23480 = 6;
      if DI23712 = 0 then DI23703=5;
              else if DI23712 in (8,9) then DI23703 = DI23712;
      if DE23111 in (2,3,4,6,8,9) then DI23248 = 3;
              else if FR23309v = 9 then DI23248 = 4;
      if DE23111 in (2,3,4,6,8,9) then HN23490 = 3;
              else if HN23309v = 8 then HN23490 = 4;
      TTHNB = coalesce(HN23301, HN23302);
              if TTHNB = . then TTHNB=5;
      S_length = round(length, 0.01);
      format HN23307 addi.
                HN23247 yesnoNev.
                HN23235 harm.
                HN23810 yesnobuy.
                HN23762 HNBbanned.
                HN23774 yesnoNon.
                HN23411 HN23415 yesnohea.
                FR23353 yesnoNS.
                BR23358 capsule.
                LM23331 agreeNS.
                CH23960 SM23942 SM23949 SM23930 yesnoQT.
                WL23221 WLharm.
```

```
                WL23313 PU23660 yesnoSMK.
                ET23634 yesnowrk.
                PR23220 DiseaLike.
                PR23313 DamLike.
                RE23255 vaperhrm.
                HN23480 friend.
                DI23703 Drinkday.
                DI23248 yesnoMar.
                HN23490 yesnoMH.
                TTHNB   tthnb.
                SB23012v TTF.
    ;
        rename ET23634 = ET23624;
        label ET23634= "People smoked at your work place";
run;


proc freq data=X15a;
        tables FR23309v*SB23012v/list missing;
        format FR23309v;
run;

/* Now re-group the X15 dataset into 3 subgroups */
/* Cig only -- Pure cig only (HNB never tried/never heard) + rest (occasional HNB
user/quitter/trier) */
/* Dual */
Data PureCig_a CigHNBtrier_a Dual_a;
        set X15a;
        if HN23350 = 3 then output Dual_a;
        else if HN23350 = 1 and HN23309v in (3:6) then output CigHNBtrier_a;
        else if HN23350 = 1 and HN23309v in (7:8) then output PureCig_a;
run;

/* Variables Christian Suggest to add
HN23158
HN23153
HN23301
HN23302
FR23118
SB23012v
SB23031
QA23101
QA23561
BI23322
BI23324
DE23211
DE23111
DE23811
Length
Device
*/
/* Use GLMselect for Group 1, the other two groups are similar */
proc GLMselect data=PureCig_a;
class
sex
AGEGRP
rSmoke
JPQregionv
HN23106 /*ever used (if heard)*/
HN23307 /*consider self addicted to HNB*/
HN23247 /*HNB brand chosen bec same as friends*/
HN23235 /*your HNB variety/flavour how harmful vs others*/
HN23810 /*bought HNB device kit with supplies*/
HN23762 /*freq seen HNB use indoors where cigs banned*/
HN23774 /*you use HNB: at bar, pub*/
HN23411 /*noticed HNB ad: email or texts*/
HN23415 /*noticed HNB ad: bars or pubs*/
FR23353 /*RYO reduces amount I smoke*/
BR23358 /*freq, crush the flavour pellet*/
LM23331 /*lights smoother on throat and chest*/
CH23960 /*used insurance-covered cessation service LQA*/
SM23949 /*used any other SSM at last QA*/
```

```sas
SM23930 /*used other NRT at last QA*/
KN23265 /*SHS causes harm to kids, elderly*/
WL23221 /*freq, cig WL stopped you from smoking*/
WL23313 /*avoid cig WL in any way*/
WL23425 /*cig WL alarm or calm you*/
ET23624 /*You smoked at workplace*/
ET23882 /*support smoking ban in restaurants*/
AD23214 /*noticed tob ads in email or texts*/
AD23217 /*noticed tob ads at bars or pubs*/
PU23660 /*to save money on cig: other*/
AD23701 /*noticed anti-smoking info at all*/
PS23328 /*society's attitude towards smoking cigs*/
IN23220 /*tobco should use plain packages for cigs*/
IN23337 /*support ban on menthol*/
PU23680 /*gov't should increase tax on cig*/
PR23220 /*smoking-related disease prob: keep smoking*/
PR23313 /*worried smoking will damage your health*/
RE23255 /*SH vapour harm vs SH cig smoke*/
HN23480 /*# HNB users of five closest friends*/
DI23712 /*how often drink alcohol*/
DI23703 /*number of alcohol drinks on typical day*/
DE23311 /*highest level of education*/
DI23248 /*partner/ spouse thinks you should quit cigs*/
HN23490 /*does spouse/ partner use HNB*/
ME23449 /*# online surveys completed recently*/
TTHNB   /* Time to the first HNB */
SB23012v /* Time to the first cig */
Device /* Device using */
;
model inM15=
sex
AGEGRP
rSmoke JPQregionv HN23106 HN23307 HN23247 HN23235 HN23810 HN23762 HN23774 HN23411
HN23415 FR23353 BR23358 LM23331 CH23960 SM23949 SM23930 KN23265 WL23221 WL23313
WL23425 ET23624 ET23882 AD23214 AD23217 PU23660 AD23701 PS23328 IN23220 IN23337
PU23680 PR23220 PR23313 RE23255 HN23480 DI23712 DI23703 DE23311 DI23248 HN23490
ME23449 TTHNB SB23012v S_Length Device
 /selection = stepwise(select=sl slentry=0.1 slstay=0.1);
run;
/* Results N = 2342 */
/*AGEGRP JPQREGIONV HN23762 SM23949 KN23265 WL23221 WL23425 PU23660 PS23328 IN23220 IN23337
HN23480 Device
Root MSE            0.35827
Dependent Mean      0.83817
R-Square            0.0775
Adj R-Sq            0.0541
AIC                 -2405.72775
AICC                -2402.51864
SBC                 -4409.96088
*/


/* CPD models */
Data X15cpd;
        merge Core(in=a) Wave1(in=b) x(in=c);
        by uniqid;
        if HN23309v = 1 then do;
                if HN23301 in (1:4) then HN23012v = 4 - HN23301;
                else if HN23301 in (8,9) then HN23012v = HN23301;
        end;
        else if HN23309v in (2,3) then do;
                if HN23302 in (1:4) then HN23012v = 4 - HN23302;
                else if HN23302 in (8,9) then HN23012v = HN23302;
        end;
        else if HN23309v in (4:8) then HN23012v = 4;
        array NU HN23506 HN23509 HN23510 HN23532 HN23514 HN23528 HN23516
                      HN23524 HN23519 HN23773 HN23771 HN23776 HN23777;
        Do over NU;
                if HN23309v in (4:8) then NU=3;
        end;
        if HN23309v in (6,7,8) then HN23566 = 6;
        array NT HN23243 HN23244 HN23249 HN23248;
```

```
Do over NT;
        if HN23309v in (7,8) then NT=3;
end;
if HN23309v in (4:8) then HN23235=4;
if HN23309v = 8 then HN23621=6;
if HN23621 = 6 then HN23622 = 6;
        else if HN23622 = . then HN23622 = 1;
if HN23309v in (4:8) then HN23810=3;
        else if HN23815 in (2,3,8,9) then HN23810=4;
if HN23309v = 8 then HN23762 = 5;
if HN23309v = 8 then HN23731 = 5;
array use HN23404 HN23406 HN23410 HN23412 HN23415 HN23419 HN23421
          HN23423 HN23702;
Do over use;
        if HN23309v = 8 then Use = 4;
end;
if HN23309v in (4:8) then HN23596=5;
if FR23326 in (1,2,8,9) then FR23333 = 4;
if FR23326 in (1,8,9) then FR23353 = 3;
if FR23309v = 1 then do;
        if SB23012 in (1:4) then SB23012v = 4 - SB23012;
        else if SB23012 in (8,9) then SB23012v = SB23012;
end;
else if FR23309v in (2,3) then do;
        if SB23027 in (1:4) then SB23012v = 4 - SB23027;
        else if SB23027 in (8,9) then SB23012v = SB23027;
end;
if QA23101 = 2 then QA23561=0;
        else if  QA23101 = 8 then QA23561=8888;
        else if  QA23101 = 9 then QA23561=9999;
if QA23561=0 then QA23671 = 3;
        else if QA23561=8888 then QA23671 = 8;
        else if QA23561=9999 then QA23671 = 9;
if QA23561=0 then SM23949 = 3;
        else if QA23561=8888 then SM23949 = 8;
        else if QA23561=9999 then SM23949 = 9;
if WL23201 = 1 then WL23211 = 1;
if ET23431 in (2,8,9) then ET23436 = 3;
if ET23531 in (2,8,9) then ET23536 = 3;
if ET23603 in (2,8,9) then ET23634 = 3;
        else if ET23603 = . then ET23634 = 4;
if ET23603 in (2,8,9) then ET23625 = 3;
        else if ET23603 = . then ET23625 = 4;
if ET23603 in (2,8,9) then ET23621 = 4;
        else if ET23603 = . then ET23621 = 5;
if BR23309v in (2,8,9) or BR23701 in (2,8,9) and PU23201 = . then PU23201=4;
        else if BR23227 in (8,9) then PU23201=BR23227;
        else if SO23225 = 76 then PU23201=5;
if HN23309v = 8 then BQ23239 = 4;
if HN23309v = 8 then HN23560 = 6;
if DI23703 not in (8,9) then DI23703=DI23703-1;
if DI23712 in (8,9) then DI23703 = DI23712;
        else if DI23712=0 then DI23703=5;
if BI23322 in (88,99) then BI23324 = BI23322;
if DE23111 in (2,3,4,6,7,8,9) then DI23245 = 3;
if DE23111 in (2,3,4,6,7,8,9) then DI23248 = 3;
if ME23449 = 1 then ME23450=0;
        else if ME23449 = 88 then ME23450 = 888;
        else if ME23449 = 99 then ME23450 = 999;
/* Add the Attrition Related Variables */
if HN23309v in (4:8) then HN23307=4;
if FR23309v = 9 then BR23358 = 6; /*non-smoker */
        else if BR23358 = . then BR23358 = 7;
if FR23309v =9 then WL23221=5;
if FR23309v =9 then PU23660=3;
if HN23309v = 8 then HN23480 = 6;
if Device = "desktop" then S_Device = 1;
        else if Device = "generic" then S_Device = 2;
        else if Device = "touch" then S_Device = 3;
format   HN23506 HN23509 HN23510 HN23532 HN23514 HN23528 HN23516
              HN23524 HN23519 HN23773 HN23771 HN23776 HN23777 yesnonon.
```

```
                          HN23566 satisfy.
                          HN23243 HN23244 HN23249 HN23248 yesnotri.
                          HN23235 Flaharm.
                          HN23621 HN23622 WLhnb.
                          HN23810 yesnobuy.
                          HN23762 smkingban.
                          HN23731 hnbrule.
                          HN23404 HN23406 HN23410 HN23412 HN23415 HN23419 HN23421
                          HN23423 yesnouse.
                          HN23702 banhnb.
                          HN23596 hnbworry.
                          FR23333 FMuse.
                          FR23353 FMonly.
                          SB23012v TTF.
                          QA23671 howquit.
                          SM23949 yesnoqt.
                          ET23436 ET23536 yesnogo.
                          ET23634 ET23625 yesnowrk.
                          ET23621 workinside.
                          PU23201 buying.
                          BQ23239 ledQT.
                          HN23560 hnbopin.
                          DI23703 drinkam.
                          DI23245 DI23248 yesnoMar.
                          HN23012v TTFhnb.
                          HN23307 addi.
                          BR23358 capsule.
                          WL23221 WLharm.
                          PU23660 yesnoSMK.
                          HN23480 friend.
                          S_Device Device.
      ;
            if a and b and c and inM15=1 then output X15cpd; /* 3143 */
      run;


      Data PureCig_b CigHNBtrier_b Dual_b;
            set X15cpd;
            if HN23350 = 3 then output Dual_b;
            else if HN23350 = 1 and HN23309v in (3:6) then output CigHNBtrier_b;
            else if HN23350 = 1 and HN23309v in (7:8) then output PureCig_b;
      run;

      /* CPD Model Selection */

       ****************************************************************************
       *********************** Group 1   ****************************************

       ********* Adding the attrition related variables **********
       *************************************************************;
       /* Grp1: AGEGRP JPQREGIONV HN23762 SM23949 KN23265 WL23221 WL23425 PU23660 PS23328 IN23220
              IN23337 HN23480 Device */
      proc glmselect data=PureCig_b
            seed=483 plots(stepAxis=number)=(criterionPanel ASEPlot);
            partition fraction(validate=0.2);
      class sex FR23309V HN23103 HN23309V HN23621 HN23622 HN23762 HN23731
       HN23404 HN23406 HN23410 HN23412 HN23415 HN23419 HN23421 HN23423 HN23702
       FR23333 FR23353 PS23243 LM23705 LM23703 SB23012v SB23031 QA23671 SM23949 KN23221 KN23265 KN23264
       KN23251 WL23201 WL23211 WL23221 WL23421 WL23427 WL23545 ET23221 ET23421 ET23436 ET23536 ET23621
       ET23634 ET23625 ET23888 ET23819 ET23792 AD23202 AD23217 AD23629 AD23627 SO23225 PU23201 PU23653
       PU23657 PU23656 AD23736 BQ23111 BQ23141 BQ23201 BQ23203 BQ23211 BQ23214 BQ23239 PS23215 PS23223
       PS23328 IN23220 IN23336 IN23327 PU23680 PR23101 BQ23301 NC23302V DI23301 HN23560 DI23712 DI23703
       DI23443 BI23322 BI23324 DE23211 DE23311 DI23245 DI23248
       AGEGRP JPQREGIONV WL23425 PU23660 IN23337 HN23480 Device;
       model FR23245v =  sex FR23309V HN23103 HN23309V
       HN23519 HN23566 HN23621
       HN23622 HN23762 HN23731 HN23404 HN23406 HN23410 HN23412
       HN23415 HN23419 HN23421 HN23423 HN23702 FR23333 FR23353 PS23243 LM23705 LM23703 SB23012v
       SB23031 QA23561 QA23671 SM23949 KN23221 KN23265 KN23264 KN23251 WL23201 WL23211 WL23221 WL23421
       WL23427 WL23545 ET23221 ET23421 ET23436 ET23536 ET23621 ET23634 ET23625 ET23888 ET23819 ET23792
       AD23202 AD23217 AD23629 AD23627 SO23225 PU23201 PU23653 PU23657 PU23656 AD23736 BQ23111 BQ23141
       BQ23201 BQ23203 BQ23211 BQ23214 BQ23239 PS23215 PS23223 PS23328 IN23220 IN23336 IN23327 PU23680
```

```
      PR23101 BQ23301 NC23302V DI23301 HN23560 DI23712 DI23703 DI23443 BI23322 BI23324 DE23211 DE23311
   DI23245 DI23248 ME23450
        AGEGRP JPQREGIONV WL23425 PU23660 IN23337 HN23480 Device
        Sex*AgeGrp    Sex*FR23309v Sex*SB23012v
        AgeGrp*FR23309v AgeGrp*SB23012v
        FR23309v*SB23012v
        HN23012v*SB23012v
 /*/selection = stepwise(select=sl slentry=0.05 slstay=0.05);*/
 /selection = stepwise(select=sl slentry=0.05 slstay=0.05 choose = cv)
                           stats    = press
                cvMethod  = random(5)
                cvDetails = all
                           hierarchy=single stb
                           /*details = all*/;
 run;


/* New Results for GRP1 after adding the interaction term:
sex FR23309V HN23731 SB23012v SB23031 KN23251 ET23536 ET23621
PU23201 BQ23141 IN23336 DE23211 AGEGRP HN23480

Root MSE            6.49781
Dependent Mean      14.56716
R-Square            0.4295
Adj R-Sq            0.4047
AIC                 7374.44412
AICC                7380.44412
PRESS               67652
SBC                 6178.55627
ASE (Train)         40.44059
ASE (Validate)      46.90451
CV PRESS            68468
 */


************ Now impute for Grp1 ******************
*************************************************;

Proc MI data= Grp1_PureCig nimpute=5 out=impute_PureCig seed = 2018
        /*2018 is just a random number used for replicate the results*/
        /*round=0.1    min=0   Max=100*/;
        class  sex FR23309V HN23731 SB23012v SB23031 KN23251 ET23536 ET23621
                  PU23201 BQ23141 IN23336 DE23211 AGEGRP HN23480;

        FCS regpmm(FR23245v = sex FR23309V HN23731 SB23012v SB23031 KN23251 ET23536 ET23621
                               PU23201 BQ23141 IN23336 DE23211 AGEGRP HN23480);

        var sex FR23309V HN23731 SB23012v SB23031 KN23251 ET23536 ET23621
             PU23201 BQ23141 IN23336 DE23211 AGEGRP HN23480 FR23245v;
run;

/* Check Grp1 imputation */
Proc means data=impute_PureCig n min max mean median;
        *where FR23309v = 1;
        by _imputation_;
        var FR23245v;
run;

Proc freq data=impute_PureCig;
        where FR23309v = 1;
        by _imputation_;
        Table FR23245v;
run;

Proc freq data=Grp1_PureCig;
        where FR23309v = 1;
        Table FR23245v;
run;

Data new_PureCig;
        set impute_PureCig;;
        if FR23245v_ori = . then imputed = FR23245v;
run;
```

```
Proc univariate data= impute_PureCig;
      var FR23245v_ori imputed;
      by _imputation_;
      histogram /normal midpoints = 2 6 10 14 18 22 26 30 34 38 42 46 50;
run;
```