



ITC Kenya Wave 1

Technical Report

June 2015



ITC Kenya Research Team

Dr. Lawrence Ikamari (Co-Principal Investigator, University of Nairobi, Population Studies and Research Institute [PSRI])

Dr. Jane Rahedi Ong'ang'o (Co-Principal Investigator, Kenya Medical Research Institute [KEMRI])

Mr. Vincent Kimosop (Institute for Legislative Affairs [ILA])

Dr. William Maina (Division of Non-Communicable Diseases, Ministry of Public Health and Sanitation)

Dr. Patrick Waihenya Mwangi (Division of Non-Communicable Diseases, Ministry of Public Health and Sanitation)

Mr. Hillary Kipruto Kipchumba (Consultant)

Dr. Richard Mutuku (University of Nairobi, Population Studies and Research Institute [PSRI])

ITC Kenya International Research Team (University of Waterloo, Canada)

Professor Geoffrey T. Fong (Principal Investigator, Department of Psychology)

Professor Mary E. Thompson (Principal Investigator, Department of Statistics and Actuarial Science)

Dr. Christian Boudreau (Co-Investigator, Department of Statistics and Actuarial Science)

Dr. Anne C.K. Quah (Research Scientist, Department of Psychology)

Dr. Susan Kaai (Project Manager, Department of Psychology)

Ms. Grace Li (Data Analyst, Department of Psychology)

Mr. Pete Driezen (Propel Centre for Population Health)

Ms. Genevieve Sansone (Student Project Manager, Department of Psychology)

Ms. Annika Green (Student Project Manager, Department of Psychology)

Funding

- Canadian Institutes of Health Research (CIHR) – 115016

Suggested Citation

ITC Project. (2015, June). *ITC Kenya Technical Report Wave 1*. University of Waterloo, Waterloo, Ontario, Canada; University of Nairobi, Ministry of Public Health and Sanitation, Kenya Medical Research Institute, Kenya.

Table of Contents

Table of Contents.....	3
List of Tables	6
List of Figures	6
1 Introduction.....	7
1.1 Background.....	7
1.2 Main Objectives	8
1.3 The Research Team	10
1.4 ITC Kenya Survey Timelines	10
2 ITC Kenya Wave 1 Sampling Design	12
2.1 Sampling Design.....	12
2.2 Sampling Frame	12
2.3 Allocation of EAs.....	15
2.4 Sampling of Households and Enumeration.....	16
2.5 Sampling of Individuals within a Household.....	17
2.6 Interview Procedures	17
2.7 Pilot Testing of Questionnaires	17
3 Household Enumeration	18
3.1 Importance of Enumeration.....	18
3.2 Definition of a Homestead/Compound	18
3.3 Definition of a Structure	18
3.4 Definition of a Private Residential Unit.....	19
3.5 Definition of a Household.....	19
3.6 Definition of a Household Head	19
3.7 Definition of a Household Member.....	20
3.8 Sampling at the Household Level	20
3.8.1 Sampling Tobacco Users	20
3.8.2 Sampling Non-users of Tobacco	20
3.9 Mapping, Listing and Random Selection of Households and Individuals....	21
3.9.1 Instructions Given to Interviewers: Map	21

3.9.2	Mapping Households.....	22
3.9.3	Structure Numbering and Household Listing.....	22
3.9.4	Filling the Household Listing Form	23
3.9.5	Visiting Selected Households.....	24
3.9.6	How to Sample 4 Tobacco-users	25
3.9.7	How to Sample 1 Non-user of Tobacco.....	27
4	Weight Construction for ITC Kenya Wave 1 Survey.....	31
4.1	General Comments about Weight Construction.....	31
5	Disposition Codes	32
5.1	Disposition Codes for Household Outcome Codes.....	32
5.2	Disposition Codes for Individual Outcome Codes.....	32
5.3	Outcome Rates for ITC Kenya Wave 1 Survey.....	33
6	Data Collection Activity.....	34
6.1	Data Collection	34
6.2	Survey Process.....	34
6.3	Survey Types.....	34
6.4	Content of Survey	36
6.5	Information Letters and Consent Forms	37
6.6	Languages of Survey Questionnaire.....	37
6.7	Collecting Empty Tobacco Packs	37
6.8	Remuneration	38
6.9	Sample Size and Representation Tables.....	38
7	Monitoring and Quality Control.....	40
7.1	Management of Fieldwork Teams.....	40
7.2	Interviewer Training Workshop	42
7.3	Interviewer Aids	43
7.4	Monitoring & Quality Assurance.....	43
7.5	Data Quality Control	44
7.6	Handling Special Situations	45
7.6.1	Private Interviews.....	45

7.6.2	A Proxy Interview	45
7.6.3	Respondent was Unavailable	45
7.6.4	Substitution	45
	Appendix A: ITC Kenya Training Workshop Agenda	46
	Appendix B: Mapping and Listing Guidelines.....	49
	Appendix C: Random Table	53
	Appendix D: Manikin Flashcards.....	56
	Appendix E: Tobacco Products.....	57
	Appendix F: Flow Charts.....	58
	Appendix G: Sampling Design and Weight Construction	71

List of Tables

Table 1: Overview of the Stratified 6–Stage Sampling Plan.....	13
Table 2: Sample Design for ITC Kenya Wave 1 Survey	16
Table 3: A Sample of the Household Listing Form.....	23
Table 4: A Sample of the Random Table	25
Table 5. Selection of Non-users by Province	27
Table 6: A Sample of the Random Table showing Gender Selection	29
Table 7: A Summary of Survey Sections by Type of Questionnaire	35
Table 8: Total Sample of Respondents by Smoking Status and Gender	38
Table 9: Description of Sample by Smokeless Tobacco Users status	39
Table 10: Types of Tobacco Users by Province	39
Table A1: ITC Kenya Training Workshop Agenda	46
Table B1: Example of how to select a non-user of tobacco in a household	49
Table C1: Random Order Sequence Table for Household Selection (Sample)	53
Table F1: Steps to Follow Before Approaching Households.....	67
Table F2: Eight Steps to Follow When Conducting Individual Interviews	69

List of Figures

Figure 1: ITC Kenya Survey Timeline	11
Figure 2: Map of Kenya Showing the 8 Provinces in 2009.....	13
Figure 3: Districts of Kenya in 2009	14
Figure 4: Example of Structure Numbering and Household Labelling.....	23
Figure 5: Management of ITC Kenya Project	40

1 Introduction

1.1 Background

The International Tobacco Control Policy Evaluation Project (the ITC Project) is a multi-country prospective cohort study designed to measure the psychosocial and behavioral impact of key policies of the World Health Organization Framework Convention on Tobacco Control (WHO FCTC). The ITC Project is a large global research initiative that currently involves more than 20 countries inhabited by over 50% of the world's population, 60% of the world's smokers, and 70% of the world's tobacco users. The ITC Project has about 100 leading tobacco researchers and is led by Professor Geoffrey T. Fong at the University of Waterloo. In 2012, two African countries (Zambia and Kenya) joined the ITC Project.

On June 25, 2004, Kenya ratified the Framework Convention on Tobacco Control (FCTC);

- In 2004, smoking was banned in restaurants, healthcare and educational facilities, universities, government facilities, public transit, indoor offices, other indoor workplaces, bars and pubs.
- In August 2007, the Kenyan government passed the Tobacco Control Act, which was assented into law on September 27, 2007. The law came into force in July 2008 and it included:
 - Display of 14 rotating text-only warnings on 30% of the front and 50% of the back of the package in Kiswahili and English.
 - Ban on misleading/deceptive packaging but does not specifically ban terms such as “light”, “mild”, and “low tar”.
 - The sale of 10 or fewer cigarettes was prohibited.
 - Ban on all advertisement of tobacco products on any medium of communication.
 - Ban on all sponsorship and smoking in public places except for designated smoking areas.
 - Ban on the sale of tobacco products to youth under 18.
 - Required the disclosure of tar, nicotine, and other toxic constituents. It is limited to the disclosure of contents and not their quantities.
 - Ban on the sale of tobacco products in automatic vending machines.
- Between 2007 and 2009, the government increased taxation of tobacco products to 5% of retail price. In 2011, the tax system on cigarettes was simplified, i.e., changed from tiered specific tax to an ad valorem tax.
- Additionally, Kenya developed the National Tobacco Control Action Plan (NTCAP) 2010-2015, which outlines the key directions for tobacco control in Kenya over a five-year period.

- On 5th December, 2014, the following new “Tobacco Control Regulations, 2014” were tabled in parliament to be implemented six months later (5th June, 2015):
 - 15 new pictorial warnings that cover 30% of the front and 50% of the back of the pack.
 - Ban on smoking in cars with children.
 - Ban on smoking on streets/walkways near public places.
 - Certification of compliance required for designated smoking areas (DSAs).
 - Requirement for manufacturers and importers to disclose product ingredients and sales/production data.
 - Restriction on public authority-tobacco industry interactions.
 - Measures to prevent industry interference.
 - Requirement for manufacturers or importers to pay 2% of value of products manufactured or imported to Tobacco Control Fund annually.

However, enforcement of these tobacco control laws is weak due to tobacco industry interference.

To evaluate the effect of the FCTC, the ITC Project is conducting parallel prospective cohort surveys with adult smokers in 22 countries – Canada, United States, Australia, United Kingdom, Ireland, Thailand, Malaysia, South Korea, China, New Zealand, Mexico, Uruguay, Germany, France, the Netherlands, Bangladesh, Brazil, Mauritius, Bhutan, India, Zambia, and Kenya.

1.2 Main Objectives

The objectives of the ITC Kenya Survey are:

- 1) To examine the prevalence and patterns of tobacco use behaviour in Kenya. The survey will also provide information about tobacco users’ knowledge, beliefs, attitudes and opinions about using tobacco.**

The ITC Kenya Survey provides multidimensional estimates of prevalence and patterns of tobacco use among the Kenyan population. It describes the population’s consumption patterns and quitting behaviour, as well as its knowledge, beliefs, and attitudes about tobacco use. Specifically, the survey investigates the population’s shift from traditional tobacco products (in the form of bidis, kreteks, and smokeless) to cigarettes.

- 2) To examine the impact of specific tobacco control policies that have been, or will be, implemented in Kenya, on tobacco use and tobacco-related behaviour among tobacco users in Kenya.**

The ITC Kenya Survey evaluates the impact of tobacco control policies in the following areas of the FCTC:

- Health warning labels and package descriptors
- Smoke-free legislation
- Pricing and taxation of tobacco products as well as the prevalence of compensatory behaviours that may offset the impact of taxation (e.g., cheaper purchasing options, smuggling)
- Education and support for cessation
- Tobacco advertising and promotion

Findings from the ITC Kenya Survey will provide a detailed picture of the current tobacco control policy landscape in Kenya, including the beliefs, attitudes and behaviours of cigarette and pipe smokers and non-smokers, following the July 2008 implementation of the WHO Framework Convention on Tobacco Control (FCTC). Of particular importance in Kenya is the linkage between tobacco control and the poverty of the tobacco users.

3) To compare the psychosocial and behavioural effects of national-level tobacco control policies and programs in Kenya with findings from the other 21 ITC countries.

The ITC Project aims to provide an evidence base to guide policies enacted under the WHO FCTC, and to systematically evaluate the effectiveness of these legislative efforts. All ITC Surveys are developed using the same conceptual framework and methods, and the survey questions are designed to be identical or functionally equivalent in order to allow strong comparisons across ITC countries. The evaluation studies conducted from the ITC Surveys take advantage of natural experiments created when an ITC country implements a policy: changes in policy-relevant variables in that country from pre- to post-policy survey waves are compared to those of other ITC countries where that policy has not changed. This research design provides high levels of internal validity, allowing more confident judgments regarding the possible causal impact of policy.

4) To suggest changes to current government tobacco policies

Recommendations to strengthen the current tobacco policies are made based on existing and derived survey information. The aim is to optimize the effects of tobacco control policies with regard to situational and individual difference moderators: (a) demographic variables; (b) personality variables (e.g., time perspective); (c) environmental context (e.g., number of peers/family members who smoke); and (d) the individual's smoking history (e.g., past quit attempts, smoking intensity, and quitting smoking).

1.3 The Research Team

The ITC Kenyan Survey was conducted by team members from the University of Nairobi, Population Studies and Research Institute (PSRI). The research team is collaborating with an international team of researchers at the University of Waterloo, Canada.

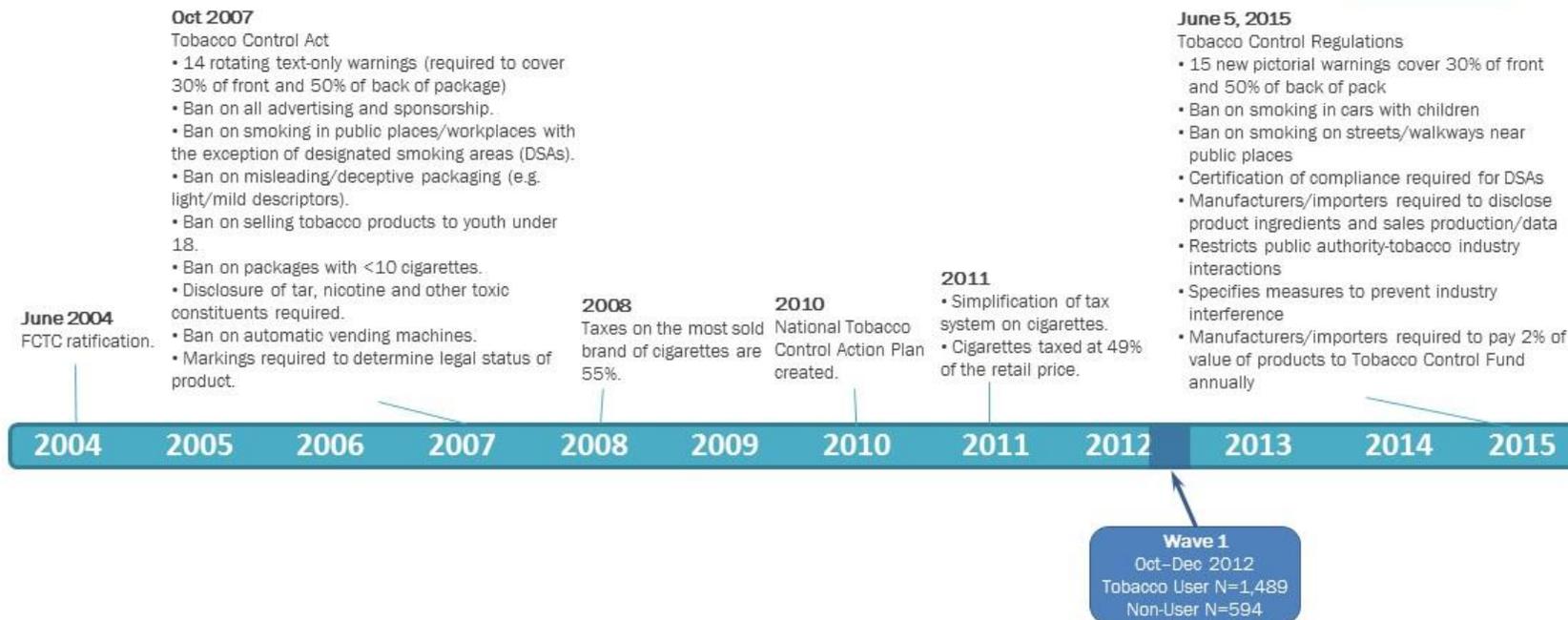
1.4 ITC Kenya Survey Timelines

The ITC Survey is a longitudinal cohort study. Therefore, the respondents who participated in Wave 1 will be re-contacted in Wave 2 to answer the follow-up survey. In Wave 1, approximately 1,500 tobacco users and 600 non-users of tobacco aged 15 and older were randomly selected for participation. For Wave 2, these respondents will be recontacted, and for those that cannot be recontacted, the sample will be replenished to retain the same approximate numbers of tobacco users and non-users of tobacco. Figure 1 shows the timeline of the ITC Kenya Project. The timing of Wave 2 is not yet determined.

Figure 1: ITC Kenya Survey Timeline

KENYA

Timeline of Tobacco Control Policies and ITC Surveys



Survey Mode: Face-to-Face (F2F)
Respondent Types: Tobacco User, Non-User

Updated March 2015

2 ITC Kenya Wave 1 Sampling Design

2.1 Sampling Design

The ITC Kenya Survey was designed to be a nationally representative sample of approximately 1,500 tobacco users and 600 non-users of tobacco aged 15 years or older selected through a multi-stage clustered sampling design. Specifically, the design was stratified by province and sampled a total of 148 clusters/enumeration areas (EA), allocated to the provinces in numbers proportional to population size. Design calculations were based on 2009 Kenya Population Census data. Within each EA, it was intended that 10 tobacco users and 4 non-users of tobacco would be interviewed. It was estimated that interviewers would have to enumerate approximately 70 households in each EA to attain 10 tobacco users. Those that currently smoked cigarettes or bidis, or used any form of smokeless tobacco at least once a month were considered to be tobacco users. An overview of the stratified multi-stage sampling plan is described below.

2.2 Sampling Frame

The sampling frame for wave 1 of the ITC Kenya Survey is the 2009 Kenya Population & Housing Census (KPHC) conducted by the Kenya National Bureau of Statistics (KNBS). According to that frame, the population of Kenya was first divided into 8 provinces; see Figure 2. Each province was then divided into districts (or wilaya) (see Figure 3), with a grand total of 158 districts. Each district was then further divided into divisions (or taarafa), which in turn were divided into locations (or mtaa) and then sub-locations (or mtaa mdogo). Finally, each of the over 7,000 sub-locations was divided into enumeration areas (EAs). These EAs consist on average of about 100 households, but their sizes varied quite a lot.

Our original sampling design consisted in stratifying the population, and then sampling districts (stage I), locations (stage II), sub-locations (stage III), EAs (stage IV), households (stage V) and finally individual respondents (stage VI). However, several locations contained few sub-locations; yielding selection probabilities close to 1 at stage III. For other locations, stage III selection probabilities would have been much smaller, and such a scenario would have resulted in sampling weights that would have been highly variable; thus decreasing precision. To avoid this, locations that contained too few sub-locations were pooled together or pooled with larger locations. This pooling yielded what we have called super-locations. In other words, super-locations are an artificial level between divisions and locations that we created. The same issue arose with sub-locations containing too few EAs, and a few sub-locations were thus pooled together. See Appendix G for a detailed report regarding how the sampling was done.

Figure 2: Map of Kenya showing the 8 Provinces in 2009

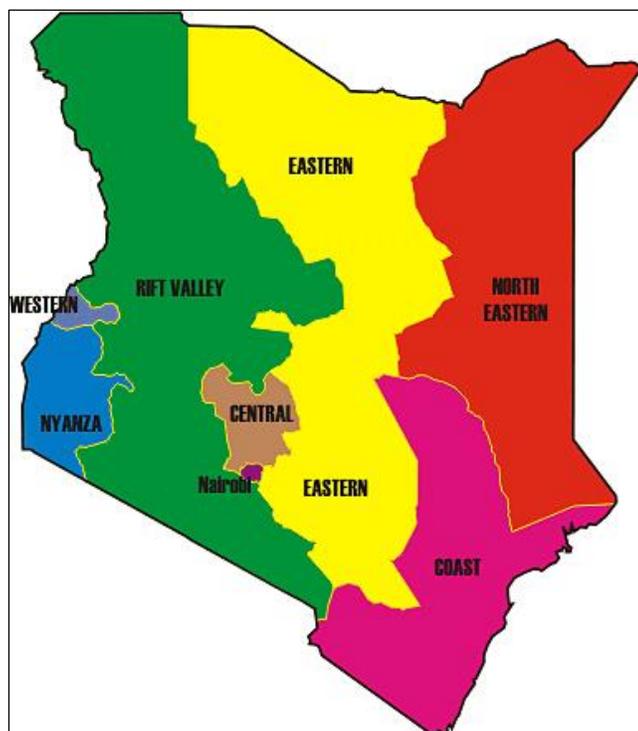
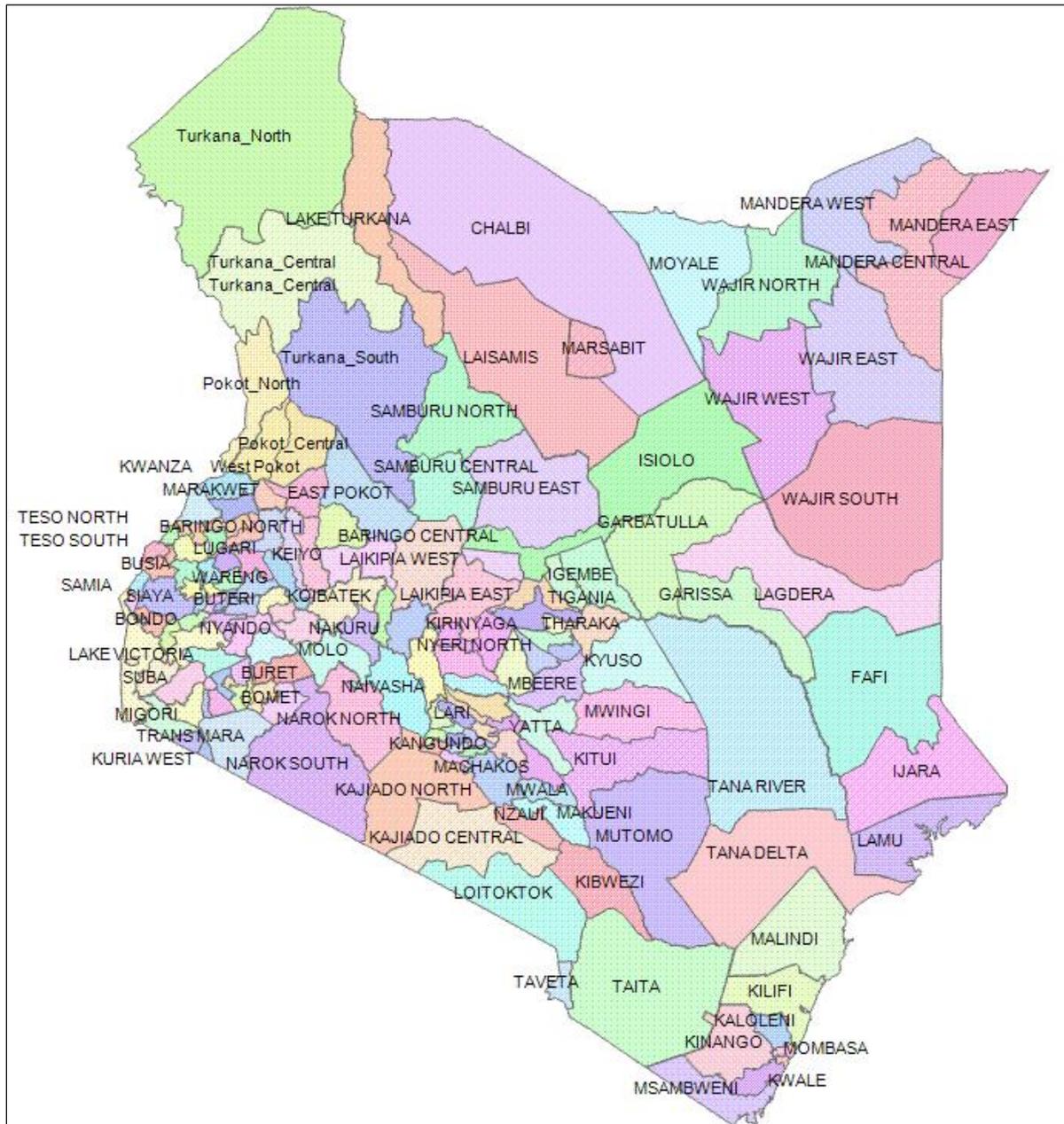


Table 1: Overview of the Stratified 6–Stage Sampling Plan

Hierarchy	Explanation
Strata	8 Provinces yielded 12 Strata
Stage 1	Sample 1-5 districts within the 12 strata (n=21)
Stage 2	Sampled 1-2 super-locations within each of 21 districts (n=37)
Stage 3	Sampled 2 sub-locations within each of 36 super-locations, while 3 sub-locations were sampled via PPS in the remaining super-location (n=75)
Stage 4	Sampled 2-3 EAs per sub-location (n=148)
Stage 5	Every household within each sampled EA was assigned a unique ID number. Those ID numbers were then randomly ordered within EA, and households were visited by interviewers following the random ordering
Stage 6	Sampled up to 4 tobacco users and 1 non-user per enumerated household. This was repeated until the quota of the EA was met i.e., 10 tobacco users and 4 non-users of tobacco

Figure 3: Districts of Kenya in 2009



There were special considerations that were made in the sampling design. They are described below:

1. Because the North Eastern Province borders Somalia, there are various safety concerns. Additionally, there are other difficulties regarding interviewing respondents in this province. In the light of this, instead of randomly sampling from the 4 districts in that province, it was decided to sample the district of Garissa with probability one (i.e., purposive sampling); see Patton(1990) for more information on purposive sampling. Note that Garissa is one of the two largest districts of that province.
2. The districts of Migori, Kuria West and Kuria East in the province of Nyanza are Kenya's main tobacco-growing regions. It was thus decided to sample one of those three districts. To this end, the province of Nyanza was divided into two strata: the first stratum consisted of the districts of Migori, Kuria West and Kuria East, whereas the second stratum consisted of the 18 other districts of Nyanza; see Table 1 in Appendix G.
3. The districts of Mombasa, Eldoret West, Eldoret East, Wareng, Busia, Teso North and Teso South are of particular interest because of potential tobacco smuggling. It was thus decided to sample 3 of those 4 districts, one per province. The district of Mombasa is the only such district in the Coast province, and it was thus sampled with probability one. As was done for the province of Nyanza, the provinces Western and Rift Valley were each divided into two strata; see Table 2 in Appendix G.

2.3 Allocation of EAs

Two to three EAs were randomly sampled with probability proportional to population size in each of the 75 sub-locations that were selected at stage three. Refer to Table 2.

Table 2: Sample Design for ITC Kenya Wave 1 Survey

Province	Sampled Districts	Sampled EAs/Clusters (n)
Central	Murang'a South	8
Central	Murang'a North	8
Coast	Kilindini	8
Coast	Mombasa	4
Eastern	Kibwezi	8
Eastern	Machakos	8
Eastern	Meru South	8
Nairobi	Nairobi East	8
Nairobi	Nairobi North	8
Nyanza	Kisii Central	8
Nyanza	Migori	4
Nyanza	Nyando	8
Rift Valley	Baringo	8
Rift Valley	Eldoret West	4
Rift Valley	Kericho	8
Rift Valley	Naivasha	8
Rift Valley	Narok North	8
Western	Bungoma South	8
Western	Lugari	8
Western	Teso	4
North Eastern	Garissa	4
TOTAL	21 DISTRICTS	148 EAs

2.4 Sampling of Households and Enumeration

The interviewer teams would go to the selected clusters also known as enumeration areas (EAs). If a list of the dwelling units of an EA was not already available, they would make a new list following written guidelines from ITC (see details in Section 3.9). They would visit the dwelling units on the list in random order, and at each dwelling unit, once contact was made, enumerate the household (i.e., enumeration included listing all the household members and their age, sex, tobacco use, and relationship to head of household), and select individuals for interview, if any were eligible. They would conduct the interviews then or at a later visit. They would stop enumerating households in the EA when the target numbers of interviews for the EA were reached. Enumeration is important for estimating prevalence, which could be done with small bias even if not all of the expected 10,500 households were enumerated, as long as households were approached in random order and the enumeration data were carefully recorded even when there was no one eligible in the households.

2.5 Sampling of Individuals within a Household

In any enumerated household all tobacco users up to a maximum of 4 (randomly selected if necessary) were to be interviewed. If there were more than 4 users in a household, the interviewers sampled all female users and selected male users at random until 4 users were sampled. This procedure was meant to increase the potential number of female users in the sample. Smoking is much less prevalent among females as compared to males. Sampling was to continue until the cluster quota of users (10 users) was reached. In every 4th (Central and Eastern Provinces), 5th (Nairobi, Coast, Rift Valley, and North-Eastern Provinces), 7th (Nyanza and western Provinces) enumerated household, up to one randomly selected non-user (male or female) was to be interviewed until the quota of non-users (4 non-users) for the cluster was met. If a 4th (or 5th or 7th, as appropriate) household did not contain any users, interviewers were still expected to sample a non-user. Among non-users, an alternation procedure was designed to sample an equal number of males and females. Details are given in section 3.9.

2.6 Interview Procedures

Data were collected through the 'face-to-face' interview method. The questionnaires were adapted by the ITC Kenya and Waterloo teams to ensure that they were relevant to the Kenyan context. They were also translated to Kiswahili (national language of Kenya) by a Kenyan translator and verified by another Kenyan translation consultant. Participants gave their informed consent before commencing an interview. The surveys took approximately 60 minutes for tobacco users and about 30 minutes for non-users of tobacco to complete. The survey was administered mainly in Kiswahili. However, the English version was also provided for respondents who wished to answer the survey questionnaire in English. A total of 3 Regional Supervisors, 19 Field Supervisors, and 45 interviewers, 1 Project Manager, and 4 data entry clerks were contracted by the University of Nairobi. Each province had a team that comprised of a Supervisor and 2-4 interviewers. Each team remained in their province until the data collection activity was completed. Data collection commenced in October 22, 2012 and was completed in December 21, 2012. A total of 6,257 households were enumerated.

2.7 Pilot Testing of Questionnaires

A pilot survey was conducted during the interviewers' training workshop on September 27, 2012 (See Appendix A for the Interviewers' Training Timetable). Interviewers had the opportunity to practice what they had learnt in the training workshop. A debriefing session was held after the pilot testing activity. Interviewers shared their experiences regarding the field logistic challenges and the actual interviewing exercise. The questionnaires were revised based on the debriefing exercise.

3 Household Enumeration

3.1 Importance of Enumeration

Household enumeration was carried out prior to administering the ITC Kenya Survey questionnaire. This activity is mainly done to:

- Obtain estimates of the total number of tobacco users among males and females, i.e., the estimates of tobacco use prevalence.
- Assess the presence/absence of children in each tobacco use/never or past tobacco-use household.
- Select survey respondents from the filled enumeration forms.
- Draw upon this information to inform the sample in subsequent waves.
- Accurately record the information in the longitudinal structure of the survey data linking the recontact sample across waves.
- Replenish samples for future waves.

The enumeration form was used to obtain the gender, age, and tobacco use status (cigarette or bidi or pipe or smokeless users) of every respondent aged 15 years and above. This information was used to select eligible household members for the interviews. A maximum of 4 attempts were made to enumerate each household. The time required to complete the Household Enumeration Form was 10-15 minutes.

3.2 Definition of a Homestead/Compound

A homestead is an isolated compound with one or more structures, and may be inhabited by one or more households. In some cases, homesteads are surrounded by fences, hedges, walls, etc. However, a boy's quarters, garage, kitchen, etc., is sometimes part of a homestead whether or not they are surrounded by a fence/wall. A homestead contains a hut or a group of huts (KNBS, 2009).

3.3 Definition of a Structure

A structure is a building that is used for the purpose of residential, business or any other activity. Eligible structures were those constituted of units used for residential purposes. Buildings used for businesses or any other activities were not eligible for this study. In the rural areas, residential structures found outside or within a homestead/compound were eligible. A structure could contain one or more residential units, e.g., a structure could be a storied building or other building containing more than one residential unit (KNBS, 2009).

3.4 Definition of a Private Residential Unit

- A private residential unit is a dwelling (with cooking facility) that is considered to be the usual place of residence for at least one of the persons living there.
- A private residential unit may be attached to a business, such as in the case of a business operating out of the home.
- The following were NOT private places of residence:
 - Addresses that are businesses only (the employees don't live there).
 - Institutions (such as schools, hospitals, nursing homes or prisons).

3.5 Definition of a Household

A household consists of the people living in a private residential unit, and may consist of the following:

- One person living alone.
- A family sharing the same residential unit.
- A family together with a roomer/boarder or employee who lives in the same residential unit.
- A group of people who are not related but share the same residence and cooking facilities or living expenses.
- Persons who reside in the same compound (i.e., the persons in the household may reside under the same roof or several roofs in the same compound), who are answerable to the same head (a person who is recognized to have authority) and have the same cooking arrangement (KNBS, 2009).
- In polygamous marriage, if the wives lived in separate dwelling units and had separate cooking arrangements, the living arrangement for each wife was treated as a separate household. Each wife with other persons who lived with her were also considered as a separate household. The husband was listed in the household that interviewers found him on the day of the data collection. But if the wives cooked in the same compound, and were answerable to one head of the household then they were all treated as one "household" (KNBS, 2009).
- It is customary in many parts of Kenya for teenage boys to live in separate quarters between circumcision and marriage, while continuing to take their meals with their parents. They are considered to be a member of their parents' households (KNBS, 2009).

3.6 Definition of a Household Head

This is the most responsible/respected member of the household who makes key decisions of the household on a daily basis and whose authority is recognized by all

members of the household. It could be a father or mother, or any other responsible member of the household depending on the status of the household (KNBS, 2009).

3.7 Definition of a Household Member

In this study, individuals who were considered members of a household were:

- Those who lived in the residential unit all year round
- Those whose permanent address was the residential unit, but were studying in another place and would only come home during school holidays
- Those whose permanent address was the residential unit, but were working abroad or in another city and would only come home once in awhile

If a member was absent for the whole period of time when fieldwork was being conducted, but still came home at least once a year, s/he could be enumerated but would not be selected for an interview. Visitors (relations or friends) who were there for a short visit or a short term stay were not considered members of the household.

3.8 Sampling at the Household Level

3.8.1 Sampling Tobacco Users

In any enumerated household, all tobacco users up to a maximum of 4 (randomly selected) were interviewed. If there were more than 4 users in a household, all female users were sampled first, and then male users were selected at random until the 4 tobacco users' maximum target was achieved. This procedure was meant to obtain as many female users in the sample as possible since smoking is much less prevalent among females than among males in Kenya. Sampling continued until the EA quota of tobacco users (10 users) was reached.

3.8.2 Sampling Non-users of Tobacco

Since it was estimated that interviewers would have to enumerate 70 households in each EA to attain 10 tobacco users, the sampling frequency of every 4th (Central and Eastern Provinces), 5th (Nairobi, Coast, Rift Valley, and North-Eastern Provinces), 7th (Nyanza and western Provinces) households was expected to be sufficient to recruit the desired quota of 4 non-users of tobacco (see more details in the mapping and listing guidelines in Appendix B). In every 4th, 5th, and 7th enumerated household, one randomly selected non-user (male or female) was interviewed until the quota of non-users (4 non-users) for the EA was met. Gender of the first non-user did not matter. After the first non-user was selected, alternating female and male non-users were selected. If a 4th, 5th and 7th household did not contain a non-user of tobacco of

the right gender, no non-user would be sampled. If a 4th, 5th and 7th household did not contain any users, sampling a non-user of tobacco of the right gender was still required.

The goal for this sampling strategy was to have an approximately equal number of female and male non-users of tobacco. This could not be achieved without the alternation strategy because a majority of non-users are female and because male non-users were expected to be absent (e.g., working away from home for long periods) or declining to be interviewed (non-response) in greater proportions.

3.9 Mapping, Listing and Random Selection of Households and Individuals

The interviewer teams were expected to go to the selected enumerated areas (EAs), draw a map of the area (if existing map was not available), and then create a list of households based on the mapping drawn. They were to visit the households on the list in random order as indicated by a random number sequence (see Random Table in Appendix C) as shown in the interviewer's field manual. At each household, once contact was made, interviewers were to enumerate the household (list all the members and their age, sex, tobacco use, relationship to head of household, etc.), and select individuals for interview, if any were eligible. They would conduct the interviews on the same or at a later visit and stop enumerating households in the EA when the target numbers of interviews for the EA were reached, i.e., 10 tobacco users and 4 non-users of tobacco.

The next sections are detailed step-by-step instructions that were given to interviewers to guide them on how to draw an EA map, how to do the household listing, and how to select households and individuals for the interviews.

3.9.1 Instructions Given to Interviewers: Map

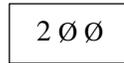
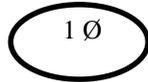
- You have been given an EA map with borders indicated.
- Your map shows the administrative unit names and codes in which the selected EA is located.
- The North Arrow: The north arrow is an arrow showing the north direction of the earth. It is provided on the map to help the user for orientation.
- Before using your map in the field, start by establishing the North direction in relation to where you are standing:
 - Stretch out your arms and stand facing the direction from where the sun rises. You are now facing east. Your left hand is pointing to the North.
 - Identify on the map a landmark to mark your current position, or one close to where you are standing (draw a star to indicate your position).

3.9.2 Mapping Households

- Ensure that your supervisor, the coordinator from Kenya National Bureau of Statistics (KNBS) office, your Enumeration Area (EA) District Statistical Officer (DSO), and the village elder are accompanying you on your first day of the EA household mapping and listing.
- Ensure that you clearly and accurately identify the EA boundaries with the help of your DSO and Supervisor.

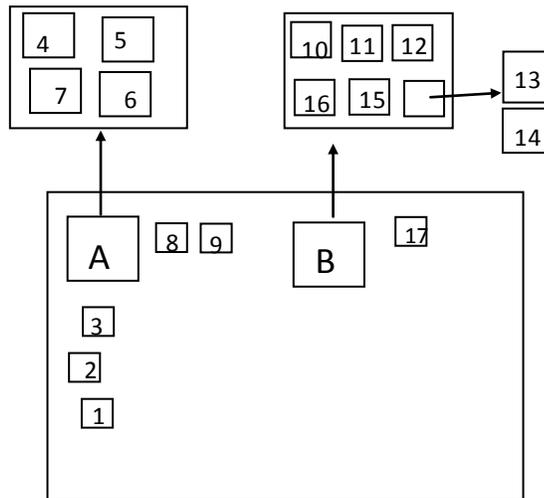
3.9.3 Structure Numbering and Household Listing

- Plot all landmarks such as churches, schools, shops, kiosks, factories, police stations, office buildings, prisons, hospitals, etc.
- Plot all households: use a circle for huts and a square for a permanent structure. Ensure that the number of the household is written inside the circle or square.



- Zero should be crossed to distinguish it from a circle or the letter 'O'.
- Once all of the households have been drawn on the map, number the households in a clockwise direction starting from 1 up to the total number of households in the selected EA.
- You may experience several scenarios depending on the types of structures:
 - a. A single household
 - b. A single household with a number of buildings or structures
 - c. Several households in one building or structure
 - d. Gated compound with several structures and several households.
- If there are multiple households in one building (structure), label the building with a capital letter (i.e., A, B, etc.) and draw an arrow outside of the box leading to a new sketch of the households within that building
 - Draw boxes and number them to represent each of the households in the building

Figure 4: Example of Structure Numbering and Household Labelling



3.9.4 Filling the Household Listing Form

- As you label each household on your map, write the address on your household listing form. The form can be completed simultaneously with the mapping of the EA.
- **Note:** You may not need to use the whole form – only fill in as many rows as you need for the total number of households in the EA.
- The form has three columns: Household number, Address, and Notes
- The address for each household should be labelled beside the same number as on the map. For example:

Table 3: A Sample of the Household Listing Form

Household No.	Address	Notes
1	302 Street Name A	
2	304 Street Name A	
3	306 Street Name A	
...	...	
10	320 Street Name A	On the right-hand side when facing the dwelling unit; blue door
11	320 Street Name A	On the left-hand side when facing the dwelling unit; brown door
12	322 Street Name A	
...	...	

Two households (numbered as units 10 and 11) within one structure with the same address

Add notes to distinguish each household within the same structure

- The first listing labelled as 1 corresponds to the address of the household numbered as 1 on the map. Structures with one address, but multiple households, will result in the same address being listed under the multiple numbers corresponding to the household numbers labelled on the map (e.g., households numbered 10 and 11).
- Make note on the list of landmarks or other points of reference in order to distinguish each household within the dwelling unit that have the same address (see example in Table 3 in the “Notes” column).

3.9.5 Visiting Selected Households

- The households are visited in a random order using a table containing randomly generated lists of numbers (See Random Table in Appendix C). The numbers in each sequence are in a random order and correspond to the numbered households on the map. **The order of the numbers as they appear in each column of the table is the order in which the households must be visited.**
- To use the table, go to the appropriate column based on the total number of households in the EA.
 - For example (see Table 4): If the EA has 119 households, go to the column labelled “between 101 and 125 units” and move down the column.
- After visiting the household, cross out the corresponding number in the table.
- Note that if you come to a number in the column that is greater than the total number of households in the selected EA, there will be no household that corresponds to this value. Therefore, ignore this number and move on the next number in the column.

Example: For an EA with 71 households: The total number of households, 71, is between 51 and 75, so you need to go to the column labeled “Between 51 and 75 units” as shown in Table below.

Table 4: A Sample of the Random Table

Between 51 and 75 units	Between 76 and 100 units	Between 101 and 125 units	Between 126 and 150 units	...
	93	93	93	...
			141	...
24	← 24	← 24	← 24	...
39	← 39	← 39	← 39	...
				...
29	← 29	← 29	← 29	...
			150	...
...

1st household visited

2nd household visited

3rd household visited

The first number in the sequence is 24. Therefore, the household numbered as '24' on the map and dwelling units/households list is the first unit to be visited. The next number in the sequence moving down the column is 39, meaning that the second household visited is the unit numbered as '39' on the map and dwelling units/households list. As we continue down the column, we see that the next number is 29. Therefore, the unit numbered as '29' will be visited after the unit labelled '39', and so on. Remember to cross out the number after the household has been visited so that you can track which households have been visited and the next household to visit.

3.9.6 How to Sample 4 Tobacco-users

- At each household unit (visited in the order determined in the table), the whole household is to be enumerated, and up to 4 ADULT tobacco users may be selected for individual interviews (eligibility status = **YT**).
- Once contact is made, fill out the Household Enumeration Form (HEF) i.e., list all the members and their age, gender, tobacco use, relationship to head of household. Details of how to fill the HEF is shown in Appendix F, Flow chart #1 and Flowchart #3).
- Up to 4 tobacco users from the household can be sampled if they are eligible. Refer to Appendix F, Flowchart #2 (titled "Respondents Selection Process") for more detailed instructions on how to select members to be interviewed.

Example: Sampling Tobacco Users

The first household to be visited is number 24 (see Table 4 in the previous page). Contact is made and the Household Enumeration Form is filled out. Suppose after enumeration we find that household 24 is composed of 2 adult non-users of tobacco, 3 children, and 3 adult tobacco users. A maximum of 4 tobacco users can be selected to be interviewed. In this case, there are only 3 tobacco users that are eligible (i.e., 15 years of age or older, specified that they use tobacco at least monthly), therefore all 3 users are selected to be interviewed.

Suppose the area is one in which a non-user of tobacco is to be interviewed in every 5th household enumerated, where possible. Because this is the 1st household we also interview one non-user of tobacco. Using a random selection method (e.g., using the lottery method), one of the two non-users is selected for interview for this household. Interviews can be conducted in the same visit for household members available and willing to participate or can be scheduled for another time for household members not available but willing to participate.

The interviewing team may then move on to the second household unit listed in the random order table, in this case, number 39. Once contact is made and the household is enumerated, we find that the household consists of 4 individuals: 3 adult non-users of tobacco and 1 child. Since this is the 2nd household visited (i.e., not one at which a non-user of tobacco can be sampled), ONLY tobacco users are eligible to be sampled. As all of the adult members of this household are non-users of tobacco, they are ineligible to be sampled. Therefore, no members from unit 39 will be interviewed. Remember to fill in your Household Enumeration Form (HEF) and submit it to the supervisor even if none of the members in the household are eligible. This HEF has to be completed because the household was sampled. Additionally, since no member of the household was eligible the Household Survey (H) will NOT be administered to the Household Head/Key Informant.

The third household to be visited is unit 29. Once contact is made and the Household Enumeration Form is filled out, we see that the household is composed of 12 members: 3 adult female tobacco users, 4 adult male tobacco users, 2 adult non-users of tobacco, and 3 children. Therefore, there are 7 eligible tobacco users in total. The 3 female tobacco users are selected to be interviewed automatically (see Appendix F, Flowchart #2 on how to select female tobacco users before males). In order to sample a maximum of 4 tobacco users, one more tobacco user must be selected for interview. Since there are 4 male tobacco users a random selection method is used to choose 1 of these individuals (see Appendix F, Flowchart # 2). Interviews can be conducted at this time or can be scheduled for another time for individuals who are not available. Keep in mind that this is the 3rd household to be visited (i.e., not the 5th household at which a non-user can be

selected); therefore, non-users of tobacco are not sampled from this household. The interviewing team will then proceed to the next household and so on.

NOTE 1: You can make up to two attempts to contact a household. If the interviewing team cannot make contact with a household on the first attempt, record the date and time of this visit, then continue on to the next dwelling unit (see Appendix F, Flowchart #1 i.e., Form A). Make sure to return to the household where no contact was initiated at a later time or date (when someone may be present) for the second attempt to make contact, before you complete the enumeration of the EA. This completes the effort to contact this household.

NOTE 2: Every household is enumerated until a total of 10 tobacco users and 4 non-users of tobacco within that EA have been interviewed. If all 10 tobacco users have been interviewed but not all of the 4 non-users of tobacco, then you need to continue to visit and enumerate to select a non-user of tobacco.

3.9.7 How to Sample 1 Non-user of Tobacco

- **Up to one** (1) adult non-user (eligibility status = **YN**) per eligible household (every 4th, 5th, or 7th household enumerated) can be selected: someone who has never used tobacco or was a past tobacco user or is a tobacco user less often than monthly.
- Start by sampling one non-user at the **first** household on your list. After that first household, count every 4th, 5th, or 7th household to find non-users.
- For **Central and Eastern Provinces**, interview one non-user at every 4th household enumerated (e.g., 1th, 5th, 9th, 13th, 17th etc) .
- For **Nairobi, Coast, Rift Valley, and North-Eastern provinces**, interview one non-user at every 5th household enumerated (e.g., 1th, 6th, 11th, 16th, 21th etc)
- For **Nyanza and Western provinces**, interview one non-user every 7th household enumerated (e.g., 1th, 8th, 15th, 22th, 29th etc).

Table 5. Selection of Non-users by Province

Province	Non-User Selection Guideline
Central and Eastern	1 non-user at 1 st household and then every 4 th household enumerated
Nairobi, Coast, Rift Valley, and North-Eastern	1 non-user at 1 st household and then every 5 th household enumerated
Nyanza and Western	1 non-user at 1 st household and then every 7 th household enumerated

Example: Using every 5th household to select one (1) non-user of tobacco

Let's take the example of selecting one non-user in every 5th household enumerated (for Nairobi, Coast, Rift Valley, and North-Eastern provinces). By every 5th household we mean the 1st, 6th, 11th households selected (as determined by the random order table) and so forth. The first household sampled is the first household listed in the appropriate column for the selected EA in the household random table. **It is NOT the household labeled '1' on the map and the dwelling units/households list.** For example, the first household at which a non-user of tobacco can be selected is the 1st household visited, which would be unit 24 (see Table 4 in section 3.9.5). The next 5th household would be the 6th household visited from the order provided in the random table. The tobacco users are selected as described in section 3.9.6. To select the non-user of tobacco to be interviewed, follow the guidelines outlined below.

The first household sampled: If the first household visited consists of only 1 non-user of tobacco, that individual is automatically selected to be interviewed. If there is more than 1 non-user of tobacco eligible in the 1st household, a random selection method (See Appendix F, Flowchart #2 titled "selecting non-users for interview") is used to choose 1 of these individuals to be interviewed. Make note of whether this selected non-user of tobacco is male or female in the last column of the household random table, labeled as 'Male/Female/No Non-User of Tobacco'.

Every 5th household sampled after the first one (i.e., the 6th, 11th, and so forth): The next non-user of tobacco selected for interview depends on the gender of the non-user of tobacco selected at the previous eligible household visited. The objective is to alternate sampling male and female non-users of tobacco at every 5th household. For example, if the non-user of tobacco selected at the 1st household visited is female, then a male non-user of tobacco is selected at the 6th household visited, a female non-user of tobacco would then be selected at the 11th household visited, and so on (see Table 6). After selection, make note of whether the selected non-user of tobacco is male or female (should be the opposite of the previous 5th household) in the last column, labelled as 'Male/Female/No Non-User of Tobacco'. If there is no eligible non-user of tobacco at the household being enumerated, indicate this in the household random table by writing 'NNU' in the last column.

Table 6: A Sample of the Random Table showing Gender Selection

	...	Between 376 and 400 units	Between 401 and 425 units	Male/Female/ No Non-user (M/F/NNU)
1 st household visited		357	357	→ F
	...	344	344	
	...	93	93	
	...	141	141	
	...	24	24	
6 th household visited		39	39	→ M
	...	224	224	
	...	29	29	
	...	150	150	
	...	179	179	
11 th household		358	358	→ F

NOTE 1: The last column of the table (Male/Female/ No Non-User) is only for every 5th household from which a non-user of tobacco can be selected. You do not need to fill in the gender in the last column for households where only tobacco users are selected (i.e., not one of the every 5th households).

In the case that there are NO eligible non-users in the current 5th household, or if the only non-user of tobacco is of the same gender as the previous non-user of tobacco selected, then the interviewer must wait until there is a gender-eligible non-user of tobacco in the next 5th household to select a non-user of tobacco to be interviewed.

In the case where a current 5th household only contain non-users of tobacco of the wrong gender, these non-users of tobacco are not eligible and are not sampled. If a non-user of tobacco is not sampled at one of the “every 5th households”, write ‘NNU’ in the last column of the household random table.

Summary:

- Continue to visit households in your list until the quota of 10 tobacco users and 4 non-users of tobacco for that EA is filled.
- **Tobacco users:** You may select up to 4 tobacco users from every household visited (with priority to female tobacco users).
- **Non-Users of tobacco:** Select 1 non-user from the first household visited (if available). After the first household, look for a non-user at every 4th household visited (in Central and Eastern provinces); or every 5th (in Nairobi, Coast, Rift

Valley, and North Eastern provinces); or every 7th household (in Nyanza and Western provinces).

- **Household Enumeration Form (HEF):** You must complete the enumeration of each selected household on your list until your quota is filled. Even if you have filled your quota of tobacco users, continue to visit households and fill out the HEF while you search for the remaining non-users from every 4th/ 5th /7th household. You will not select any individuals for survey interviews from the households with no eligible non-user.
- **The Household (H) Survey-** Only administer the Household Survey to the Household Head or Key Informant in a household that has members who are eligible tobacco users (YT status) or Non-users (YN status)---**1 Household = 1 H Survey.**

4 Weight Construction for ITC Kenya Wave 1 Survey

4.1 General Comments about Weight Construction

As with most survey weights, the ITC Kenya weights were constructed to correct and adjust for sample misrepresentation caused by unequal sampling probabilities, frame error (i.e., under-coverage and multiplicity), and non-response as well as improving precision of estimates through the use of auxiliary information (e.g., tobacco usage prevalences). We briefly describe these key concepts of weight construction in this section, but refer the reader to Levy & Lemeshow (2008), chapter 16, for more detailed information.

At their base, sampling weights were defined as the inverse of selection probabilities, and thus adjusted for sample misrepresentation caused by unequal sampling probabilities. For example, a tobacco user residing alone had a probability of selection twice than that of a tobacco user residing with another tobacco user.

Great efforts were made to create a complete/perfect sampling frame (i.e., a frame that would include all members of the target population, without duplicate and without any erroneous inclusions). However, this is rarely achieved and, consequently, some members of the target population were not part of the sampling frame (i.e., have a 0 probability of being selected). This is referred to as frame under-coverage, and can result in non-coverage bias. To reduce non-coverage bias in the ITC Kenya Survey, post-stratification adjustments were performed on the sampling weights to ensure that, for each sex/age/province group for males tobacco users and sex/age group for females tobacco users, the totals of the sampling weights equal estimated number of such tobacco users based on the enumeration Data. Please refer to Appendix G for the full details of how the weights were constructed.

5 Disposition Codes

5.1 Disposition Codes for Household Outcome Codes

1. Could not find
2. Vacant
3. Not a household (e.g., business premise)
4. Threat to safety
5. No contact- weather condition
6. No answer- 2 attempts
7. No answer- survey period ends
8. Household refusal
9. Language barrier
10. No one capable of answering
11. Enumeration prevented for other reasons
12. Enumerated

5.2 Disposition Codes for Individual Outcome Codes

1. Missed (after 4 attempts)
2. Language barrier
3. Health/Mentally incapable
4. Proxy refusal
5. Refusal
6. Incomplete (start, break-off)
7. Completed
8. Away for the entire survey period

5.3 Outcome Rates for ITC Kenya Wave 1 Survey

Number of households attempted = 8,770

Number of households contacted = 6,418

Household contact rate = $6,418/8,770 = 73.2\%$

Number of households enumerated = 5,454

Household enumeration rate, given contact = $5,454/6,418 = 85.0\%$

Household response rate = $5,454/8,770 = 62.2\%$

Number of tobacco users identified for interview = 1,734

Number of tobacco users interviewed = 1,427

Tobacco user individual response rate, given selection = $1,427/1,734=82.3\%$

No. of eligible non-users in households from which a non-user was to be selected = 1,847

Number of non-users identified for interview = 1,000 (approximate)

Number of non-users interviewed = 571

Non-user individual response rate, given selection = $571/1000 = 57.1\%$ (approximate)

Total number of interviews is 1,998

6 Data Collection Activity

6.1 Data Collection

The period of the data collection was from October 22, 2012 to December 21, 2012. The mode of interviewing was face-to-face, using household enumeration forms to obtain information from the household. Randomly selected individuals participated in the individual surveys based on their tobacco use status.

6.2 Survey Process

In Wave 1, the ITC Survey protocol consisted of 7 main steps:

1. Household enumeration using the Household Enumeration Form
2. Administering the Household Survey to the Head or Key Informant
3. Random selection of respondents to be interviewed
4. Obtaining informed consent using the Consent Form
5. Identification of appropriate questionnaire for respondent using Screener
6. Administering the individual survey questionnaire
7. Exit and provision of token of appreciation.

6.3 Survey Types

Below are the 5 types of questionnaires and a screener that interviewers used in this survey. Interviewers used the screener to identify the correct type of questionnaire to use for each sampled respondent. The questionnaire was administered after the household enumeration was completed:

- **Household (H) questionnaire:** Only for the Head or Key Informant of the Household.
- **Smoked Tobacco (T) questionnaire:** For respondents who smoke cigarettes (factory-made only or hand-rolled only or both).
- **Smokeless Tobacco user (L) questionnaire:** For respondents who use smokeless tobacco.
- **Mixed Tobacco user (M) questionnaire:** For respondents who use both cigarettes and smokeless tobacco.
- **Non-user of tobacco (N) questionnaire:** For respondents who do not smoke cigarettes or use any smokeless tobacco products.

Screener: For every respondent to ensure that the appropriate individual survey questionnaire (T, L, M or N) was administered (see Appendix F, Flow Chart #4).

All surveys had similar characteristics. Table 7 shows a description of these characteristics.

Each survey was divided into a number of sections that were arranged in a specific order. Table 7 shows the sections that were relevant to each survey. In each survey section, the total number of questions across the different surveys was not necessarily the same.

Table 7: A Summary of Survey Sections by Type of Questionnaire

Survey Sections		Survey Type			
		M	T	L	N
SMOKELESS	PAST & PRESENT FREQUENCY	√	√	√	√
	WHEN AND WHY	√		√	
	DEPENDENCE	√		√	
	QUITTING ATTEMPTS	√		√	
	BELIEFS ABOUT QUITTING	√		√	
	BRAND CHOICE AND PURCHASE	√		√	
	PERCEIVED RISK	√		√	
	PSYCHOSOCIAL BELIEFS	√	√	√	√
	KNOWLEDGE OF HEALTH EFFECTS & TOBACCO CONSTITUENTS	√	√	√	√
	WARNING LABELS	√	√	√	√
CIGARETTE	PAST & PRESENT FREQUENCY	√	√	√	√
	WHEN AND WHY	√	√		
	DEPENDENCE	√	√		
	QUITTING ATTEMPTS	√	√		
	BELIEFS ABOUT QUITTING	√	√		
	BRAND CHOICE AND PURCHASE	√	√		
	LIGHT/MILD	√	√	√	√
	PERCEIVED RISK	√	√		
	PSYCHOSOCIAL BELIEFS	√	√	√	√
	KNOWLEDGE OF HEALTH EFFECTS & TOBACCO CONSTITUENTS	√	√	√	√
WARNING LABELS	√	√	√	√	

Survey Sections		Survey Type			
		M	T	L	N
OTHER SMOKED TOBACCO	PAST & PRESENT FREQUENCY	√	√		√
PIPE	PAST & PRESENT FREQUENCY	√	√	√	√
	BRAND CHOICE AND PURCHASE	√	√		
BIDI	PAST & PRESENT FREQUENCY	√	√	√	√
ENVIRONMENTAL TOBACCO SMOKE		√	√	√	√
CESSATION HELP		√	√	√	
ANTI-TOBACCO CAMPAIGNS		√	√	√	√
TOBACCO PROMOTION		√	√	√	√
TOBACCO INDUSTRY		√	√	√	√
MODERATORS		√	√	√	√
DEMOGRAPHICS		√	√	√	√
SURVEY CLOSING		√	√	√	√

6.4 Content of Survey

The research design focuses on how individuals respond to policies and how they change over time. Below is a general description of the main constructs assessed in the ITC surveys:

- Demographic variables: These include questions to assess: gender, age, ethnicity, education, number of smokers in the household; smoker's state of health, religion, socioeconomic status.
- Proximal variables: These include measures assessing awareness of a policy (e.g., of warning labels, cessation assistance, advertising and promotion) and, where relevant, cognitive processing as a result of exposure to the policy (e.g., thinking about health warnings).
- Distal variables: Questions assessing distal variables include those that test psychosocial theories (e.g., the theory of planned behaviour: attitudes, subjective norms, perceived behavioural control/self-efficacy), risk perceptions, quit intentions, and other relevant measures. In addition, the survey included

questions that measure smokers' self-exempting beliefs, that is, those that many smokers hold that may help to sustain their smoking behaviour.

- Moderator variables: Questions about moderator variables include items assessing perceived time perspective (i.e., the tendency for individuals to think about the long-term versus short-term consequences of their actions, which is a predictor of smoking behaviour) and stress in addition to the demographic variables listed above. It also examines background variables such as country, region, and community size.
- Tobacco use behaviour variables: Standard questions assessing tobacco use behaviours developed by the WHO were utilized. There are questions that measure a variety of aspects of tobacco use behaviour including usual brand, quit intentions, and other smoking-relevant constructs.

6.5 Information Letters and Consent Forms

The interviewer first allowed the respondent to read the information letter about the research project (or the interviewer used the information letter as a guide when orally explaining the study details to respondents with low literacy) before conducting the interview. Once the project was explained, the respondent completed two copies of the Consent Form as proof of consent to participate in the study. The interviewer gave the respondent the information letter and a copy of the signed (or thumb stamped if the respondent could not sign) consent form. The second copy of the consent form and the screener were attached to the completed survey questionnaire for that respondent.

6.6 Languages of Survey Questionnaire

The survey was administered in Kiswahili which is the national language. Survey questionnaires in English were also provided for respondents who preferred to answer the survey questionnaire in English. A highly trained and experienced translator was hired to translate the English questionnaires into Kiswahili as per the ITC standard translation protocol used in the other ITC countries. The checking exercise was also done as stipulated in the same protocol.

6.7 Collecting Empty Tobacco Packs

During the data collection exercise, empty tobacco packs provided voluntarily by respondents who smoked cigarettes or bidis were collected by the interviewers and were subsequently handed to the Field Supervisors for storage.

6.8 Remuneration

For Wave 1, tobacco user respondents received a gift (T-Shirt) that cost approximately 600 Kenya Shillings (KSH) as a token of appreciation for their participation.

6.9 Sample Size and Representation Tables

A total of 5,454 households were enumerated at Wave 1. Individual interviews were conducted in 1,776 of those, yielding a sample of 1,427 tobacco users and of 571 non-users; the total sample consisted of 1,998 respondents. Table 8 shows a breakdown of the sample by smoking status and gender. Table 9 shows a breakdown of the sample by smokeless tobacco use status.

Table 8: Total Sample of Respondents by Smoking Status and Gender

Sample Area	Smoked tobacco only (N)		Smokeless only (N)		Mixed user (N)		Non-user of tobacco (N)	
	Male	Female	Male	Female	Male	Female	Male	Female
Central	144	5	1	6	0	0	33	30
Coast	86	13	9	3	2	0	23	23
Eastern	158	8	18	42	3	0	52	44
Nairobi	125	17	11	0	6	1	30	32
Nyanza	136	12	12	16	4	0	37	38
Rift Valley	197	17	68	66	3	2	71	64
Western	113	19	48	17	1	0	41	38
North Eastern	32	0	4	0	2	0	7	8
Subtotal A	991	91	171	150	21	3	294	277
Subtotal B	1,082		321		24		571	
Grand Total	1,998							

Table 9: Description of Sample by Smokeless Tobacco Users Status

Sample Area (provinces)	Uses Smokeless Tobacco		Does not use smokeless tobacco	
	Male	Female	Male	Female
	N	N	N	N
Central	1	6	177	35
Coast	11	3	109	36
Eastern	21	42	210	52
Nairobi	17	1	155	49
Nyanza	16	16	173	50
Rift Valley	71	68	268	81
Western	49	17	154	57
North Eastern	6	0	39	8
Subtotal A	192	153	1285	368
Subtotal B	345		1,653	
Grand Total	1,998			

Table 10: Types of Tobacco Users by Province

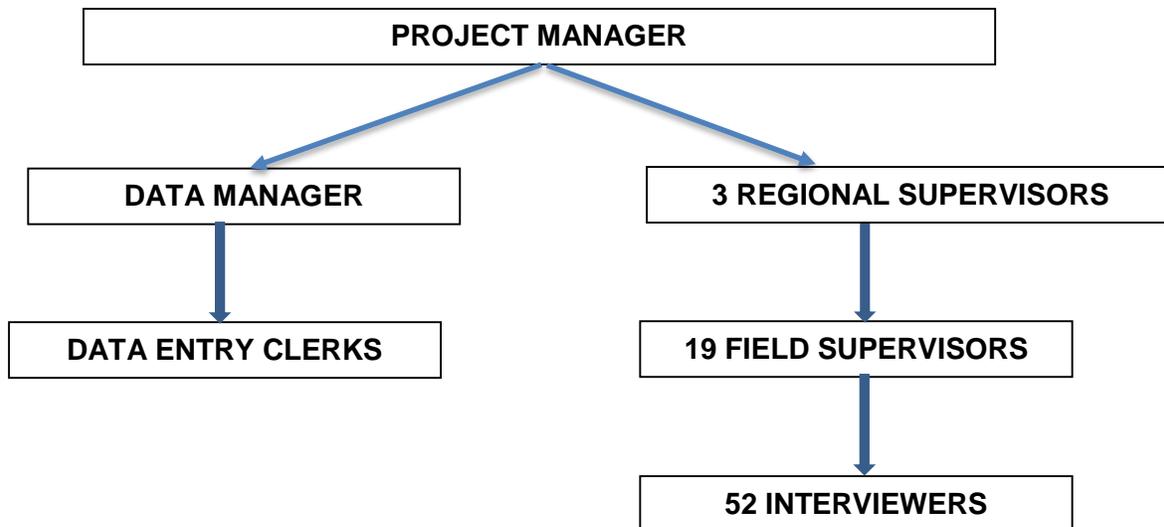
Sample Area (provinces)	Smoker				Non-Smoker			
	Does not use smokeless tobacco (smoked tobacco only)		Uses smokeless tobacco (mixed user)		Does not use smokeless tobacco (non-user)		Uses smokeless tobacco (smokeless only)	
	Male	Female	Male	Female	Male	Female	Male	Female
	N	N	N	N	N	N	N	N
Central	144	5	0	0	33	30	1	6
Coast	86	13	2	0	23	23	9	3
Eastern	158	8	3	0	52	44	18	42
Nairobi	125	17	6	1	30	32	11	0
Nyanza	136	12	4	0	37	38	12	16
Rift Valley	197	17	3	2	71	64	68	66
Western	113	19	1	0	41	38	48	17
North Eastern	32	0	2	0	7	8	4	0
Subtotal A	991	91	21	3	294	277	171	150
Subtotal B	1,082		24		571		321	
Subtotal C	1,106				892			
Grand Total	1,998							

7 Monitoring and Quality Control

7.1 Management of Fieldwork Teams

The project fieldwork team consisted of four levels of management as shown below,

Figure 5: Management of ITC Kenya Project



Their responsibilities included the following activities:

Project Manager:

- Overseeing all aspects of the survey fieldwork & data quality
- Completing administrative responsibilities of the project
- Providing necessary guidance to team members
- Communicating with the University of Waterloo (UW) on every aspect of the project
- Overseeing data entry process and check data entry errors
- Checking and reporting fieldwork progress to UW

Regional Supervisors:

- Overseeing all aspects of the survey fieldwork to ensure that the fieldwork is executed as planned in the region
- Providing necessary guidance to team members in the respective regions and communicating any challenges to the Program Manager

- Monitoring the field Supervisors to ensure that they checked and signed the completed questionnaires
- Compiling field reports and submitting them to the Project Manager

Data Manager:

- Collecting and checking all completed forms and surveys
- Testing that the epi-data template sent from UW was correctly done
- Overseeing the double data entry process
- Overseeing the checking, cleaning and compilation of data
- Transferring data from paper to computer (double entry), data cleaning and transfer of data to the Data Management Centre (DMC) at UW

Field supervisors:

- Informing relevant local authorities about the survey
- Identifying HHs selected from each Village/EA and ensuring all interviewers are ready with the HH list for each Village/EA
- Managing surveys and consent forms, and ensuring that interviewers had all the materials for fieldwork
- Traveling with the team and addressing any questions or concerns from the interviewers
- Checking the quality of information given by respondents and recording by the interviewers
- Communicating with the Project Manager about progress and difficulties encountered in the field
- Completing the fieldwork progress report and sending it to the Project Manager
- Handling travel arrangements and other field logistics

Interviewers:

- Obtaining consent from each survey respondent
- Interviewing each respondent using the correct survey based on the smoking status of the respondent (from the Screener)
- Reporting any problems or concerns to the field supervisor

Each province was covered by one team comprising 1 supervisor and 2-4 interviewers.

7.2 Interviewer Training Workshop

The ITC Waterloo Project Manager (PM) prepared a training manual and schedule for the ITC Kenya training workshop. The facilitators for the training included the ITC Chief Principal Investigator, Prof. Geoffrey Fong, ITC Africa Project Manager (Dr. Susan Kaai), and three other ITC Waterloo staff members, namely Dr. Anne Quah (Research Scientist) and Dr. Genevieve Sansone (Student Project Manager) and Ms. Annika Greene (Student Project Manager). The in-country Principal Co-Investigators (Dr. Lawrence Ikamari and Dr. Jane Ong'ang'o) and a staff member from the Kenya National Bureau of Statistics office in Kenya also facilitated the mapping session during the training. It was conducted on September 24-28, 2012 at the Multimedia University College, Nairobi, Kenya. The training program is shown in Appendix A. The training workshop consisted mainly of PowerPoint presentations role plays and actual fieldwork practice to pilot test the questionnaires. After the pilot testing activity, a debriefing session was held and the information was used to improve the English and translated questionnaires. Interviewers also received feedback on how to improve their interviewing skills. On the last day of training, a management meeting was held with the ITC Kenya and Waterloo team to discuss the revision of certain questions, translations, and fieldwork logistics.



7.3 Interviewer Aids

Several interviewing aids were used to facilitate the administration of questionnaires. They include the following:

- **Manikin Flashcards:** There were two questions in each of the surveys that required the aid of a flashcard when the response options were read out, to save time and to facilitate interviewing. The flashcards included pictures of little man-like figures (manikins) with bipolar degrees of emotion (from negative to positive emotions). Researchers studying emotional responses found the use of these manikins to be helpful to respondents in rating their own emotional responses. (See Manikin Flashcards in Appendix D).
- **Reference Sheets:** Interviewers were also provided with reference sheets that included pictures of different tobacco products in the Kenyan market (Appendix E) and step-by-step guidelines on how to conduct the data collection activity (see Appendix F, Flowchart # 4, 5, 6, Table F1, and Table F2).

7.4 Monitoring & Quality Assurance

To ensure the accuracy and the quality of the ITC Kenya Survey, the fieldwork was monitored in several ways.

Regional and Field Supervisors: The Field Supervisor travelled with each interviewer team and provided regular feedback to the interviewers. The Field Supervisor ensured that the survey protocol and data collection standards were being closely followed. Field Supervisors could monitor interviews and were available to address any questions or concerns from the interviewers. The Regional Supervisors monitored the quality of the fieldwork specifically to see that the Field supervisors executed their responsibilities well.

Identification Numbers: Field Supervisors were instructed to ensure that households and respondents identification numbers were being properly filled.

Checking for completeness: At the end of each day, interviewers were required to perform a self-check on the survey questionnaires they completed that day. The interviewer reviewed the whole survey to determine whether any questions had erroneously been missed or skipped. The Field Supervisor then collected all completed questionnaires from the interviewers under their supervision and conducted a thorough check to ensure that they were correctly filled. If there were gaps, the Fields Supervisor asked the interviewer to re-visit the household member and complete the questionnaire. Upon the completion of each survey in each village/ward the Field Supervisor gave the completed and checked questionnaires

to Regional Supervisors who conducted a final check before sending them to the in-country Project Manager (PM), who then gave it to the data entry Manager.

Weekly Meetings: The attendees of these in-country meetings were the Co-Investigators (Co-PIs), Project Manager (PM), Data Entry Manager (DM), Project Statistician, Project Accountant (PA), and the Regional Coordinators. The meetings were held regularly, i.e., at least once in two weeks, and were mainly attended by the whole team. However, sometimes a smaller core team comprising the Co-PIs, PM, DM, and PA held the meetings.

The objectives of the meetings were to check and monitor the fieldwork progress, deal with matters arising from the field, and monitor the budget and the expenditures of the fieldwork and the data sending process. Issues related to the change of enumeration areas where resistance was met, or any other challenges encountered in the field, were discussed in these meetings.

Progress Reports: The in-country Project Manager provided regular email updates, 2-weekly progress reports on quotas completed and a narrative report regarding the progress of the data collection exercise. He also informed the Waterloo PM about any concerns or problems that arose in the field.

7.5 Data Quality Control

The double entry of data proceeded in parallel with the data collection activity. In order to ensure the quality of the data collection process, the team used a multistage monitoring system:

- There were random visits by the in-country regional coordinators to monitor the interviewers in the field
- Enumerated households were randomly called to verify the information that the enumerators filled on the forms
- Field Supervisors cross-checked all completed enumeration and surveys daily to ensure that they were properly filled
- After data entry was completed, the Data Manager ran routine checks on the data sets and informed the in-country PM about any potential problems. When any issues arose the in-country PM contacted the supervisors using mobile phones. The Field Supervisors would then relay the message to the interviewers.

7.6 Handling Special Situations

7.6.1 Private Interviews

The standard was that the participants would be privately interviewed alone; however in some circumstances some household members insisted on being present during the interview. If such a situation arose the interviewer would only proceed with the interview if the respondent was agreeable to having the person present during the interview.

7.6.2 A Proxy Interview

A proxy interview is an interview conducted with another knowledgeable member of the household on behalf of the selected respondent. An example would be a woman answering the survey for her husband. Proxy interviews were not allowed in the ITC Kenya Survey.

7.6.3 Respondent was Unavailable

If a respondent was unavailable, an appointment time was rescheduled to a time that was convenient to the respondent. Only four such attempts were allowed.

7.6.4 Substitution

A substitution from the same household was allowed ONLY if a selected individual from the Non-User category (N status) had Individual Outcome Code 2 (Language barrier) or Code 3 (Health/mentally incapable) or Code 8 (Away from the household for the entire survey period).

Appendix A: ITC Kenya Training Workshop Agenda

University of Nairobi, Population Studies and Research Institute (PSRI), Nairobi

Date: 24-28 September, 2012

Venue: Multimedia University College of Kenya

Note taker: Annika Green

Table A1: ITC Kenya Training Workshop Agenda

Day 1 – Monday 24 September, 2012		
Time	Description	Presenter
8:30 - 9:00am	Registration	In-country Project Manager (ICPM)
9:00 – 9:30am	Welcome <ul style="list-style-type: none"> • Overview of Workshop Agenda and Goals • Introduction of Team Members • Ice Breaking Activity 	Lawrence Ikamari
9:30 – 9:45am	Overview of Tobacco Control in Kenya <ul style="list-style-type: none"> • Importance of survey 	Jane Ong'ang'o
9:45-10:00am	Introduction to the ITC Project	Geoffrey Fong
10:00-10:30am	Principles of Conducting ITC Surveys and Interviewing Techniques	Susan Kaai
10:30 –10:45am	Refreshments	
10:45-12:00noon	Sampling Plan and Fieldwork Procedures Enumeration Areas Mapping	Anne Quah Lawrence Ikamari/ KBS
12:00-1:00pm	Lunch	
1:00-2:30pm	Importance of Household Enumeration and Description of Household Enumeration Form	Susan Kaai
2:30-3:30pm	Role Play for Household Enumeration (including Q & A)	All
3:30-3:45pm	Refreshments	
3:45-4:00pm	Description of Information Letter and Consent Form	Genevieve Sansone
4:00-4:30pm	Overview of Survey types and Survey Features	Susan Kaai
4:30-5:00pm	Debriefing	Lawrence Ikamari/ Jane Ong'ang'o/Anne Quah
Day 2 – Tuesday 25 September, 2012		
Time	Description	Presenter

8:30-9:15am	Description of Household (H) Survey (includes Q&A)	Susan Kaai
9:15-10:00am	Role Play of Household Survey (includes Q&A)	All
10:00-10:15am	Screeener	Genevieve Sansone
10:15-10:30am	Refreshments	
10:30-12:30	Description of Smoked Tobacco User Survey (T) (Includes Q & A)	Anne Quah
12:30-1:00pm	Lunch	
1:00-2:00pm	Role Play of Smoked Tobacco User (T) (includes Q & A)	All
2:00-3:00pm	Description of Smokeless Survey (L) (includes Q & A)	Genevieve Sansone
3:00-3:15pm	Refreshments	
3:15-4:15pm	Role Play of Smokeless Survey (L) (includes Q & A)	All
4:15-4:30pm	Description of Non-User Survey (N) (includes Q & A)	Susan Kaai
4:30-5:00pm	Debriefing	Lawrence Ikamari/ Jane Ong'ang'o/Anne Quah
Day 3 - Wednesday, 26 September, 2012		
Time	Description	Presenter
9:00 – 1:00pm	Public Event at 680-Mkutano Room	
1:00-2:15pm	Lunch/Travel to training venue	
2:15-2:30pm	Description of Mixed User Survey (M) (includes Q & A)	Susan Kaai
2:30-3:30pm	Role Play of Mixed User Survey (M) (includes Q & A)	All
3:30-3:45pm	Refreshments	
3:45-4:40pm	Description of Interviewer Scripts and Flowcharts	Anne Quah

4:40-5:25pm	Mock Interviews using Kiswahili Documents <ul style="list-style-type: none"> • HEF • Screener 	All
5:25– 5.30pm	Debriefing	Lawrence / Jane /Anne
Day 4 - Thursday, 27 September, 2012		
Time	Description	Presenter
8:30-10:30	Mock Interviews using Kiswahili Documents Surveys	All
10:30-11:30am	Fieldwork Organization and Instructions	Lawrence/ICPM/ Jane Ong'ang'o
11:30– 12:30pm	Lunch	
12:30-5:00pm	Visit to the Field to Practice Interviews	All
5:00– 5.30pm	Debriefing	Lawrence / Jane /Anne
Day 5—Friday, 28 September, 2012		
Time	Description	Presenter
8:30-10:15am	Feedback from Field Practice Interviews	All
10:15-10:30am	Refreshments	
10:30-11:30am	Feedback from Field Practice Interviews	All
11:30-12:30am	Debriefing and Wrapping of Training: <ul style="list-style-type: none"> • Feedback, Evaluation Forms, Q&A 	Lawrence Ikamari/ Jane Ong'ang'o /Anne Quah
12:30-1:30pm	Lunch	
1:30-2:30pm	Meeting with Supervisors – Fieldwork Logistics	All
2:30-3:30pm	Management Team Meeting	All
3:30-3:45pm	Refreshments	
3:45-4:45pm	Management Team Meeting	All
4:45-5:00pm	Wrap-up	Lawrence Ikamari/ Jane Ong'ang'o / Anne Quah

Appendix B: Mapping and Listing Guidelines

There are several possible scenarios that may occur in the field when selecting a non-user. This section covers several guidelines on what to do when you encounter these scenarios.

Example: Central and Eastern provinces (select a non-user from every 4th household):

Table B1: Example of how to select a non-user of tobacco in a household

1 st household	2 nd household	3 rd household		4 th household	5 th household
<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#93</div> <ul style="list-style-type: none"> • HEF • Household Survey • 1 non-user • Up to 4 tobacco users 	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#39</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users 	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#29</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users 	•	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#58</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users 	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#48</div> <ul style="list-style-type: none"> • HEF • Household Survey • 1 non-user • Up to 4 tobacco users
6 th household	7 th household	8 th household		9 th household	10 th household
<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#49</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users 	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#62</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users 	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#79</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users 	•	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#45</div> <ul style="list-style-type: none"> • HEF • Household Survey • 1 non-user • Up to 4 tobacco users 	<div style="border: 1px solid black; width: 40px; height: 25px; margin: 0 auto; text-align: center; line-height: 25px;">#74</div> <ul style="list-style-type: none"> • HEF • Household Survey • Up to 4 tobacco users

Scenario 1:

You interview a female non-user at the first household (#93).

Continue to sample tobacco users at the 2nd, 3rd, and 4th households. At the 5th household on the list (#48), look for a male non-user.

If there is no non-user or only female non-users at household #48, do not select any non-users. You must wait until the next 4th household (#45) to select a male non-user.

Scenario 2:

There are no non-users at the first household (#93).

Wait until the next 4th household (#48) to look for a non-user. You can select either a male or a female non-user.

Scenario 3:

You were unable to make contact at household #48 on your first attempt, so you don't know if there are any eligible non-users there.

Continue to enumerate and interview at households #49, 62, and 79. Before you go to the next 4th household (#45), go back to #48 and make your second attempt to enumerate. If there is still no answer, you may go to the next 4th household to try to find a non-user.

Scenario 4:

At the 2nd household (#39), there are no tobacco users.

Complete the HEF. Do not do a Household Survey because there are no eligible survey respondents in the household. Do not select any individuals for interview.

Scenario 5:

You complete your quota of 10 tobacco users when you finish the 5th household (#48), but you have not completed your quota of non-users yet.

- Continue to visit households #49, #62, #79 but only do the household enumeration (HEF). Do not select any more tobacco users. There will be no Household Surveys and no individual surveys because there are no eligible respondents in those households.
- At the next household visited (the 9th household, #45), you will try to complete the HEF AND select one non-user if available. If a non-user is eligible, you will also complete the Household Survey. No tobacco users would be selected from household #45.
- Continue this procedure until you have filled your quota of 4 non-users for the EA.

Scenario 6:

You complete your quota of 4 non-users before you have completed your quota of tobacco users.

Continue to visit and enumerate households in the random order determined by the table. You may select up to 4 tobacco users at each of these households until you complete the quota of 10 tobacco users. Do not select any more non-users from every 4th household.

Scenario 7:

You complete your quota of tobacco users but you still have an appointment with a selected tobacco user from the last household to go back to do an interview. Or, you have enumerated and selected at least one tobacco user from that household but do not actually have an appointment, and before you make contact at the individual level, you fill your quota of tobacco users for the EA.

- Even though you have already filled your quota, you should go back to that household to complete the interview at the appointed time. In this case, you will exceed your quota of tobacco users. Even if you do not have an appointment, you should still go back to interview any individuals who have already been selected.
- NOTE: You should try to avoid this situation by keeping a continuing record of the number of interviews completed in each category and the number of individuals who have been selected but not yet interviewed. This should help you keep track of how close you are to completing your quota so that you do not make more appointments than necessary (ie. if you have already identified and selected enough individuals that would reach your quota, you should not go too far ahead in the enumeration process before completing the interviews with those individuals).

Scenario 8:

You have already interviewed 8 tobacco users in an EA, so you only need two more to complete your quota. At the next household you visit, you find 3 eligible tobacco users.

You may select all three tobacco users and complete the individual interviews if they are available. In this case, you will exceed your quota of tobacco users by one.

Scenario 9:

You select a household but when you go for your first household visit, you find out the entire household is away for a funeral and will be back before enumeration in that EA is finished, but not for several days, eg. 3, 4, or 5 days.

- You may continue to visit selected households in the order provided. If you fill your quota of tobacco users before the household comes back, you may return to that household for your second household attempt, but only complete the enumeration – do not select any individuals at that household for interview. If you have not yet filled your quota of tobacco users, you may select up to 4 tobacco users from that household.
- Note: If that household had been one of the 4th households in which you could select a non-user (e.g. #48), you may continue on your list and go to the next 4th household (#45) to try to find a non-user. Be sure to follow the gender requirement - If you needed a male non-user at #48, you will try to get a male at #45 instead. Once you are able to go back to #48, you should still try to get a non-user of alternating gender if there is one eligible, and if you have not yet completed your quota of non-users for the EA. If you have already filled your

quota of non-users before going back to that household, you will not select a non-user.

Scenario 10:

You have visited all of the households selected for the EA and you have not yet fulfilled your quota of tobacco users and/or non-users.

If you have exhausted your EA, inform your supervisor and they will provide you with a neighbouring EA that you can use to complete your quota.

Scenario 11:

One of the households that has been selected to visit in the EA you are working in happens to be your own household.

If your household is selected in the household random table, you should not visit that household to do the enumeration or any survey interviews. You must cross it off the random list and skip to the next household instead.

Appendix C: Random Table

Table C1: Random Order Sequence Table for Household Selection (A 3-page sample)

76 - 100 HH unit s	101- 125 HH unit s	126- 150 HH unit s	151- 175 HH unit s	176- 200 HH unit s	201- 225 HH unit s	226- 250 HH unit s	251- 275 HH unit s	276- 300 HH unit s	301- 325 HH unit s	326- 350 HH unit s	351- 375 HH unit s	376- 400 HH unit s	401- 425 HH unit s	426- 450 HH unit s	451- 475 HH unit s	476- 500 HH unit s	Male/ Female / No Non- User (M/F/ NNU)
																484	
					216	216	216	216	216	216	216	216	216	216	216	216	
															466	466	
64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	
				185	185	185	185	185	185	185	185	185	185	185	185	185	
												399	399	399	399	399	
					220	220	220	220	220	220	220	220	220	220	220	220	
											374	374	374	374	374	374	
39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	
							254	254	254	254	254	254	254	254	254	254	
						238	238	238	238	238	238	238	238	238	238	238	
														439	439	439	
24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	
							257	257	257	257	257	257	257	257	257	257	
															468	468	
								280	280	280	280	280	280	280	280	280	
									307	307	307	307	307	307	307	307	
80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	

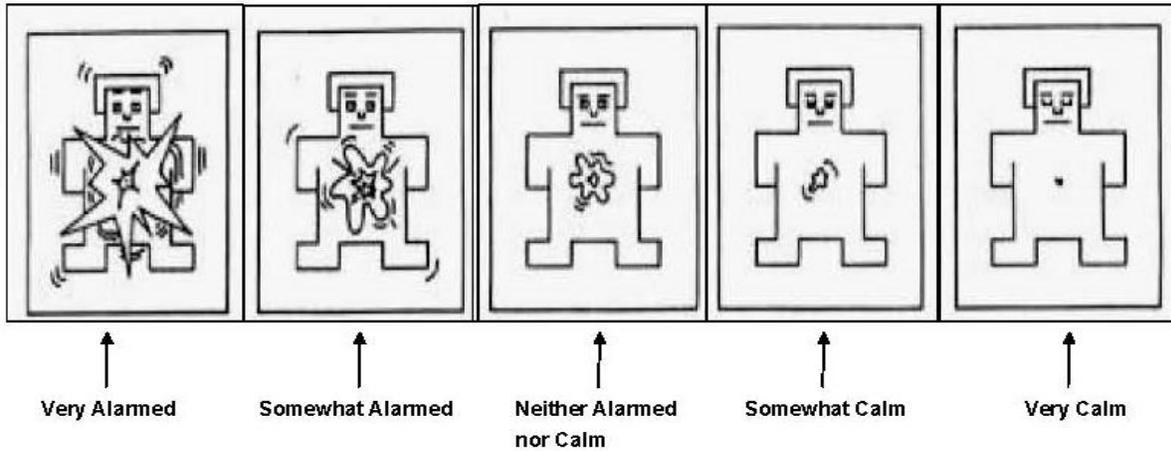
76 - 100 HH unit s	101- 125 HH unit s	126- 150 HH unit s	151- 175 HH unit s	176- 200 HH unit s	201- 225 HH unit s	226- 250 HH unit s	251- 275 HH unit s	276- 300 HH unit s	301- 325 HH unit s	326- 350 HH unit s	351- 375 HH unit s	376- 400 HH unit s	401- 425 HH unit s	426- 450 HH unit s	451- 475 HH unit s	476- 500 HH unit s	Male/ Female / No Non- User (M/F/ NNU)
											369	369	369	369	369	369	
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
															458	458	
													416	416	416	416	
27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	
									316	316	316	316	316	316	316	316	
												381	381	381	381	381	
91	91	91	91	91	91	91	91	91	91	91	91	91	91	91	91	91	
18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	
30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	
						227	227	227	227	227	227	227	227	227	227	227	
										330	330	330	330	330	330	330	
73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	
							272	272	272	272	272	272	272	272	272	272	
				181	181	181	181	181	181	181	181	181	181	181	181	181	
							273	273	273	273	273	273	273	273	273	273	
															459	459	
		145	145	145	145	145	145	145	145	145	145	145	145	145	145	145	
			164	164	164	164	164	164	164	164	164	164	164	164	164	164	
											356	356	356	356	356	356	
58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
																497	

76 - 100 HH unit s	101- 125 HH unit s	126- 150 HH unit s	151- 175 HH unit s	176- 200 HH unit s	201- 225 HH unit s	226- 250 HH unit s	251- 275 HH unit s	276- 300 HH unit s	301- 325 HH unit s	326- 350 HH unit s	351- 375 HH unit s	376- 400 HH unit s	401- 425 HH unit s	426- 450 HH unit s	451- 475 HH unit s	476- 500 HH unit s	Male/ Female / No Non- User (M/F/ NNU)
						249	249	249	249	249	249	249	249	249	249	249	
						233	233	233	233	233	233	233	233	233	233	233	
						228	228	228	228	228	228	228	228	228	228	228	
																490	
								297	297	297	297	297	297	297	297	297	
40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	
							267	267	267	267	267	267	267	267	267	267	
		130	130	130	130	130	130	130	130	130	130	130	130	130	130	130	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
											351	351	351	351	351	351	
21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	
								276	276	276	276	276	276	276	276	276	
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
														449	449	449	
											367	367	367	367	367	367	
										326	326	326	326	326	326	326	
			159	159	159	159	159	159	159	159	159	159	159	159	159	159	
36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	
					224	224	224	224	224	224	224	224	224	224	224	224	
											370	370	370	370	370	370	
								287	287	287	287	287	287	287	287	287	
														433	433	433	
																499	
													411	411	411	411	

Appendix D: Manikin Flashcards

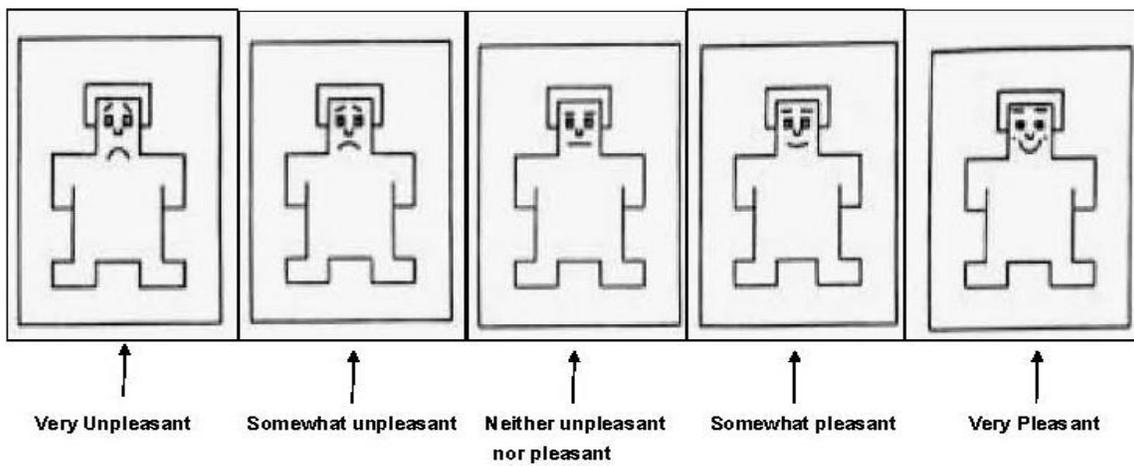
MANIKIN FLASHCARD 1

AROUSAL SCALE



MANIKIN FLASHCARD 2

VALENCE SCALE



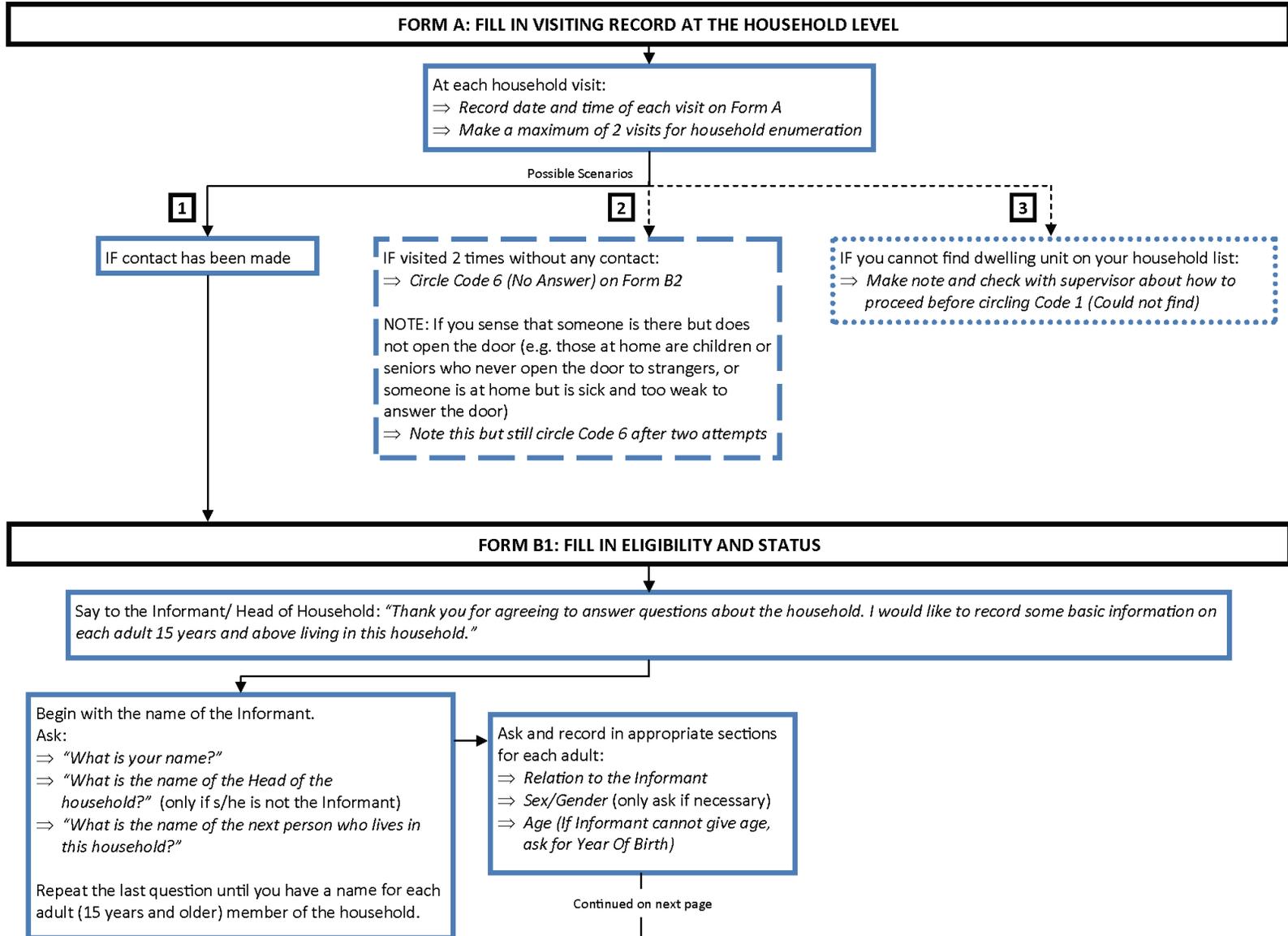
Appendix E: Tobacco Products

Types of Tobacco Products

<p>Bidi – tobacco wrapped in tendu or temburni leaves (native to India)</p>  <p>Source: http://www.gethealthyclarkcounty.org/live-tobacco-free/other-bidis-kreteks.php</p>	<p>Kretek – a cigarette made of tobacco and cloves (native to Indonesia)</p>  <p>Source: http://www.smokznews.net/wto-us-appeal-clove-cigarette-case/</p>
<p>Cheroot – a cigar with both ends clipped and left open</p> <p>Chutta – a coarsely made cheroot, typically homemade</p>  <p>Source: http://taccoon357.multiply.com/journal/item/996/The_jassic_cowboy_cigar_known_as_The_C_heroot</p>	<p>Hookah / water pipe – a pipe used to smoke shisha in which the smoke is cooled before inhalation by being drawn through water</p>  <p>Source: http://www.easterncorner.com/hookah-1-detail.htm</p>
<p>Pipe – a reusable device used for smoking tobacco, consisting of a bowl, stem, and mouthpiece</p> <p>Hookli – clay pipe (common in India)</p>  <p>Source: http://www.lotusalk.com/forums/T3/ca-smog-bottom-line-82913/</p>	<p>Cigar – a tightly rolled bundle of tobacco</p>  <p>Source: http://www.photo-dictionary.com/phrase/530/cigar.html#</p>
<p>Electronic cigarette – a battery powered device that delivers nicotine, flavor, and other chemicals</p>  <p>Source: http://www.smokedoptalk.com/2010/05/everything-about-ecigarettes/</p>	<p>Cigarillo – a short, narrow cigar</p>  <p>Source: http://www.mikescigars.com/brands/chevere-cigarillo</p>

Appendix F: Flow Charts

Flowchart 1: Steps for Completing Household Enumeration Form (Form A, B1, B2)—Continued from last page



FORM B1: FILL IN ELIGIBILITY AND STATUS

For Columns C,B and L:

Determining status

Ask the Informant the following 3 questions for each adult in the house. Begin with the Informant:

1. Do you [Does <name>] currently smoke **cigarettes** at least once a month? Yes or No or Unknown* (Record **Y** or **N** or **U** in Column C)
2. Do you [Does <name>] currently smoke **bidis** at least once a month? Yes or No or Unknown* (Record **Y** or **N** or **U** in Column B)
3. Do you [Does <name>] currently use **smokeless tobacco** at least once a month? Yes or No or Unknown* (Record **Y** or **N** or **U** in Column L)

Determining eligibility for selection:

An individual is **eligible** for random selection (Record **YT** or **YN** in Eligibility Column) if:

1. There is at least one 'Y' in the four columns (C, B, L) – the individual is **eligible as a tobacco user (Record YT)**
2. All three columns have a 'N' and the household is one of the **"1 in 4" households** (Central and Eastern provinces) or **"1 in 5" households** (Nairobi, Coast, Rift Valley, and North-Eastern provinces) or **"1 in 7" households** (Nyanza and Western provinces) from which a non-user can be selected **AND** his/her gender matches what is required from that household to meet alternating gender requirement – the individual is **eligible as a non-user of tobacco (Record YN)**

An individual is **not eligible** for random selection (Record **NN** in Eligibility Column) if:

1. All three columns have a 'N' and the household is **NOT** one of the **"1 in 4" households** (Central and Eastern provinces) or **"1 in 5" households** (Nairobi, Coast, Rift Valley, and North-Eastern provinces) or **"1 in 7" households** (Nyanza and Western provinces) from which a non-user can be selected
2. All three columns have a 'N' and the household is one of the **"1 in 4" households** (Central and Eastern provinces) or **"1 in 5" households** (Nairobi, Coast, Rift Valley, and North-Eastern provinces) or **"1 in 7" households** (Nyanza and Western provinces) from which a non-user can be selected, but his/her gender does **NOT** match the required gender for that household to meet alternating gender requirement
3. The three columns have a combination of 'U's and 'N's for status, with at least one 'U'

FORM B2: FILL IN TOBACCO USE STATUS

Adult tobacco users and non-users by gender

⇒ Record the number of male and female tobacco users and non-users, and any members with Unknown status by counting from the individuals you listed in Form B1

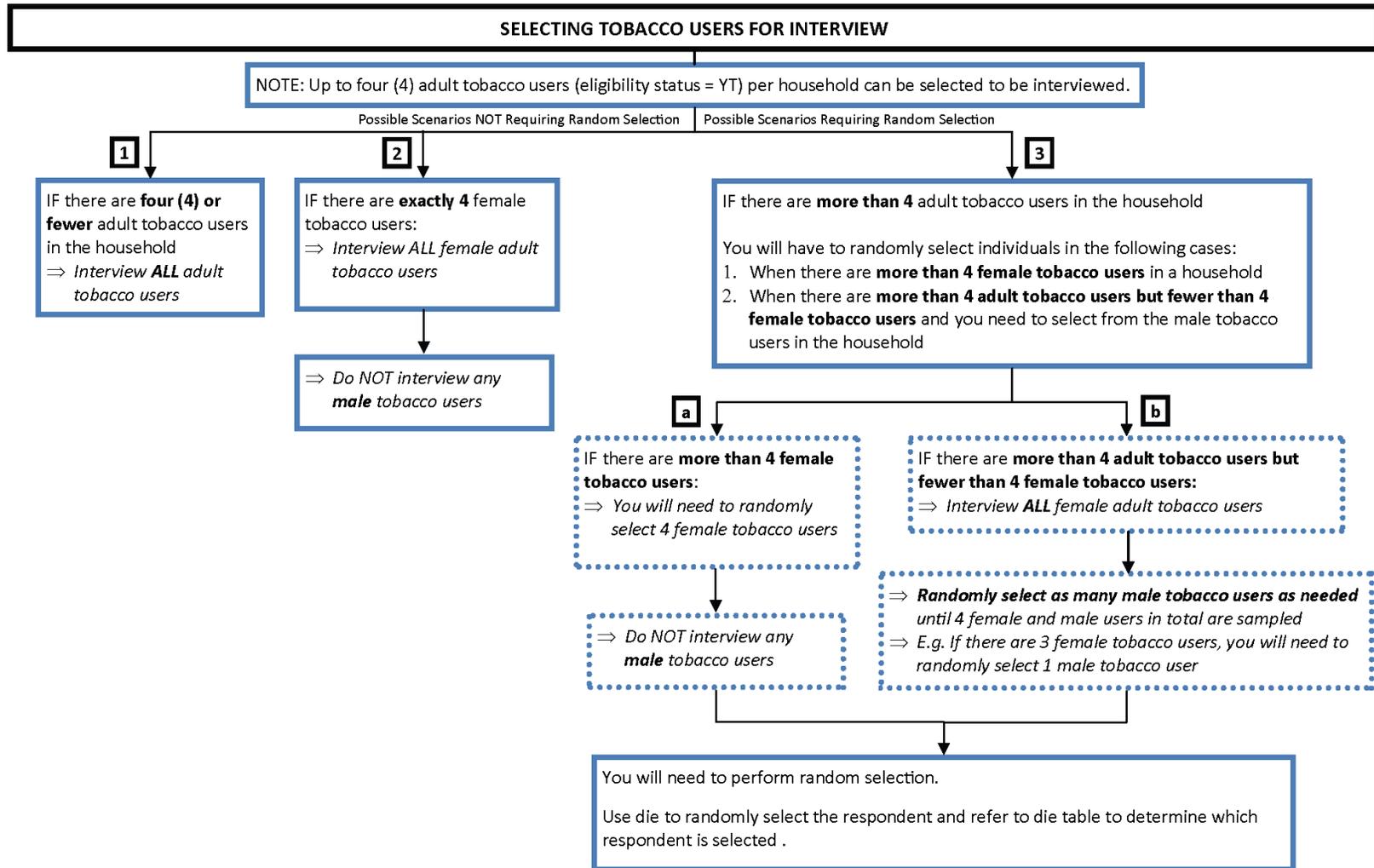
Number of children in the household

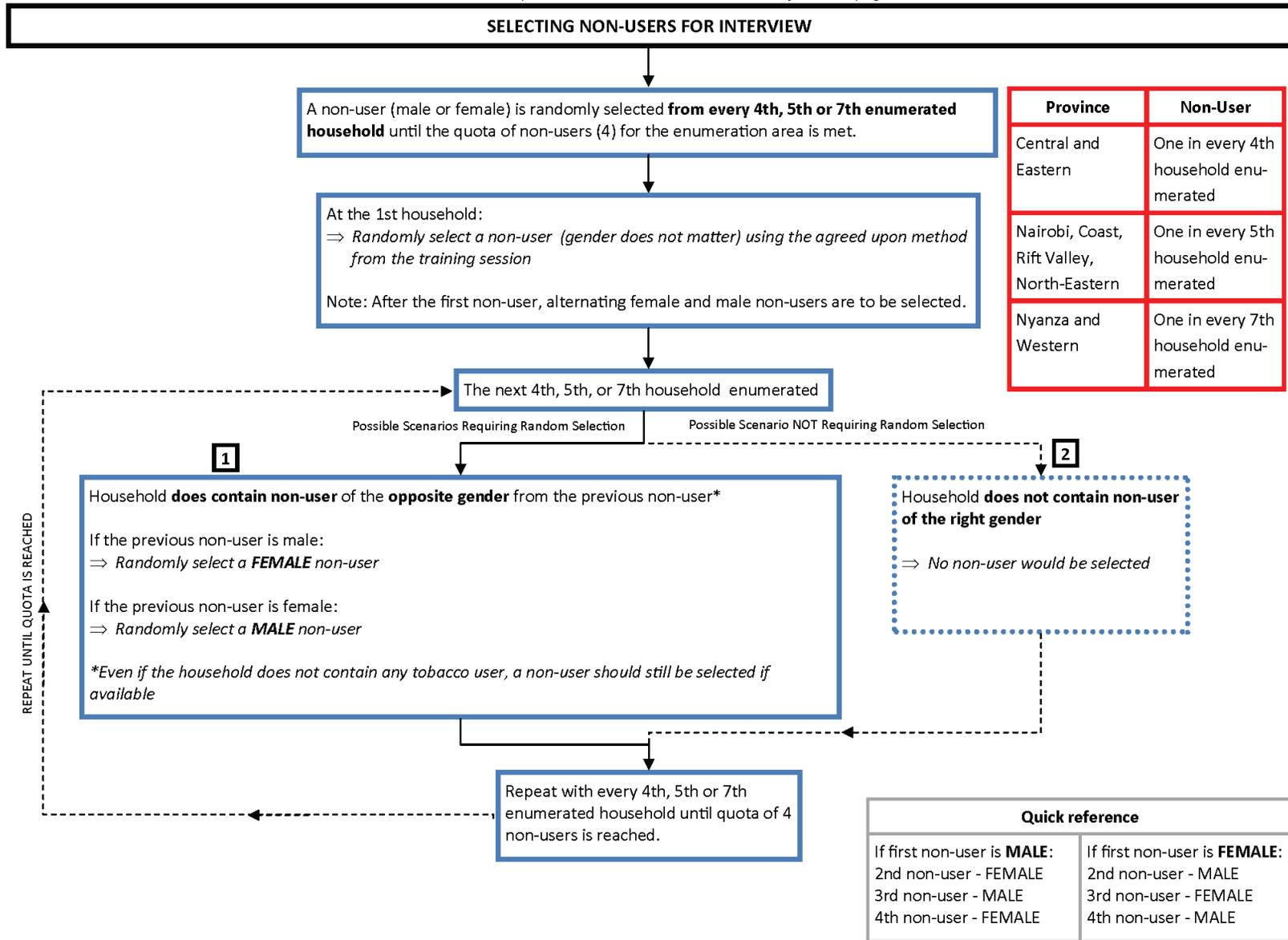
⇒ Ask and record the number of children (14 years or younger) based on age group

Final household outcome

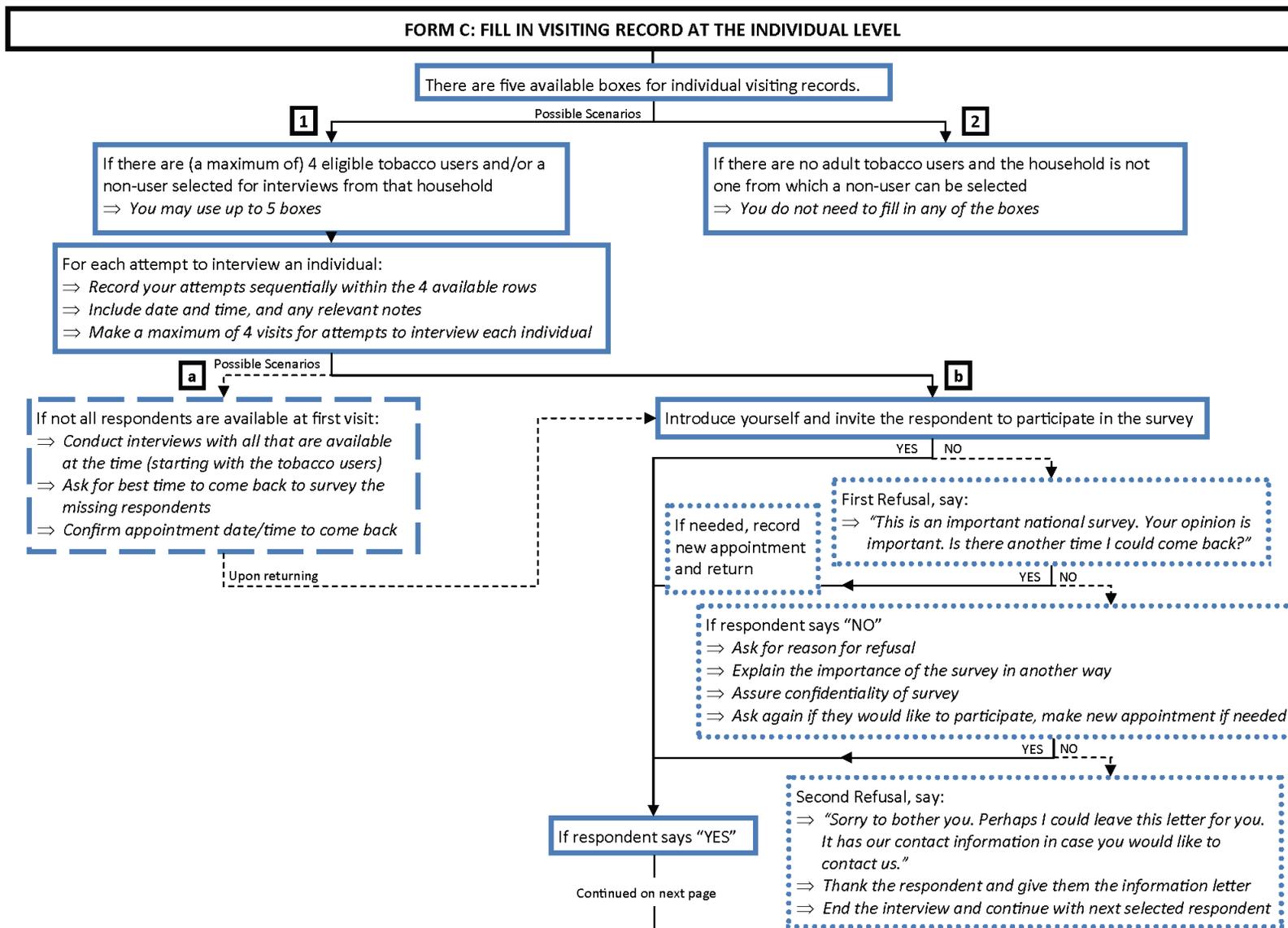
⇒ Make sure you circle one of the Final Household Outcome codes
 ⇒ Once the household is fully enumerated, circle Code 12 and continue with individual surveys

Flowchart 2: Respondent Selection Process





Flowchart 3: Steps for Completing Household Enumeration Form (Form C and D)



FORM C: FILL IN VISITING RECORD AT THE INDIVIDUAL LEVEL

⇒ Ask respondent to sign two copies of consent form

⇒ Conduct the individual survey

⇒ Thank the respondent and provide appropriate token of appreciation

After the interviews have been conducted or attempted:

- ⇒ Record appropriate Individual Outcome Code for each member of the household
- ⇒ Record Code 7 (Completed), once individual has completed the survey

Note: If the selected non-user is too intoxicated to complete the survey (after 4 attempts), use Individual Outcome Code 3 (Health Mentally Incapable).

FORM D: FILL IN ADDRESS AND CONTACT INFORMATION

After completing enumeration of each household, be sure to record:

- ⇒ Address and telephone number for the household
- ⇒ Information about contact person: Name, address, telephone number, relationship to the Informant

Flowchart 4. Screener Administration

BEFORE ADMINISTERING THE SCREENER

Before administering the screener and appropriate individual survey:

- ⇒ Determine the respondent's unique identification number (this is the 2-digit ID number from Form B1 and B2)
- ⇒ Record this unique ID number on the screener and selected survey questionnaire

ADMINISTERING THE SCREENER

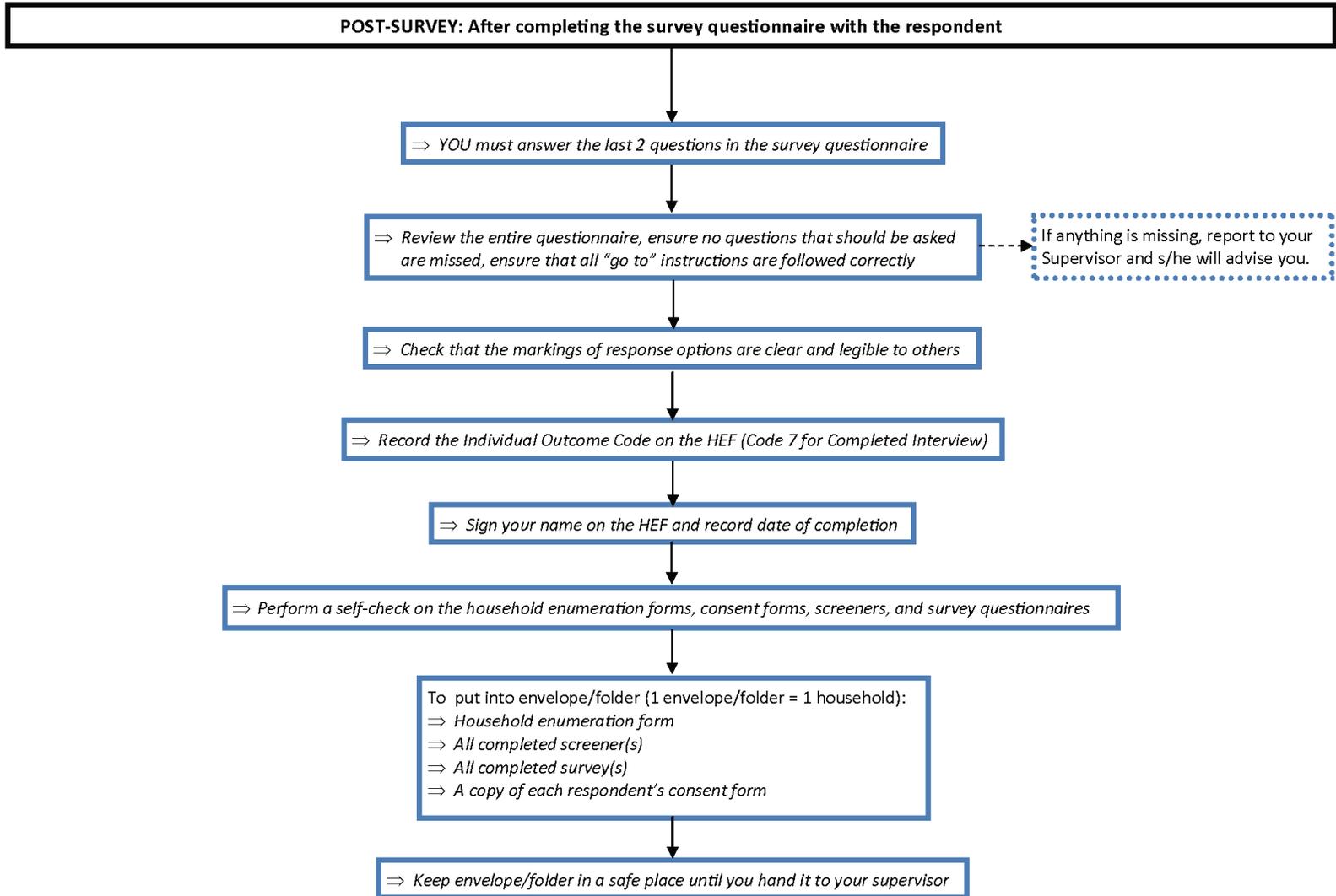
Fill in cover page of Screener

- ⇒ **Interviewer ID number**
- ⇒ Unique **household ID**: Province (2 digits), District (2 digits), Division (2 digits), sub-Location (2 digits), E.A. No. (3 digits), E.A. Status (1 digit), Structure No. (5 digits), and HH No. (3 digits). No other characters and symbols are allowed. This ID should be exactly the same as the one for the Household Enumeration Form.
- ⇒ **Individual ID** number as written in the Household Enumeration Form.
- ⇒ Fill in the **date** of survey completion and the time you start (**Start time**) the survey. When you complete the survey fill in the **End time**.

Asking the screener questions and identifying the appropriate survey questionnaire

- ⇒ Ask the respondent the **3 questions** on the cover page and **check (✓)** the correct box for each question (Yes or No).
- ⇒ In order to identify which questionnaire to use for the respondent, use the answers from the first 2 questions on the cover page to **clearly circle** the appropriate boxes in the flow chart in page 2 of the screener.
- ⇒ Once you have identified the survey questionnaire that you will use, write the survey type (C, L, M, or N) on the cover of the screener
- ⇒ Fill in your **Supervisor's name**. Your Supervisor will add the date after reviewing the form for completeness and accuracy.

Flowchart 5. Post-Survey



Flowchart 6: Overview of Steps for Each Household

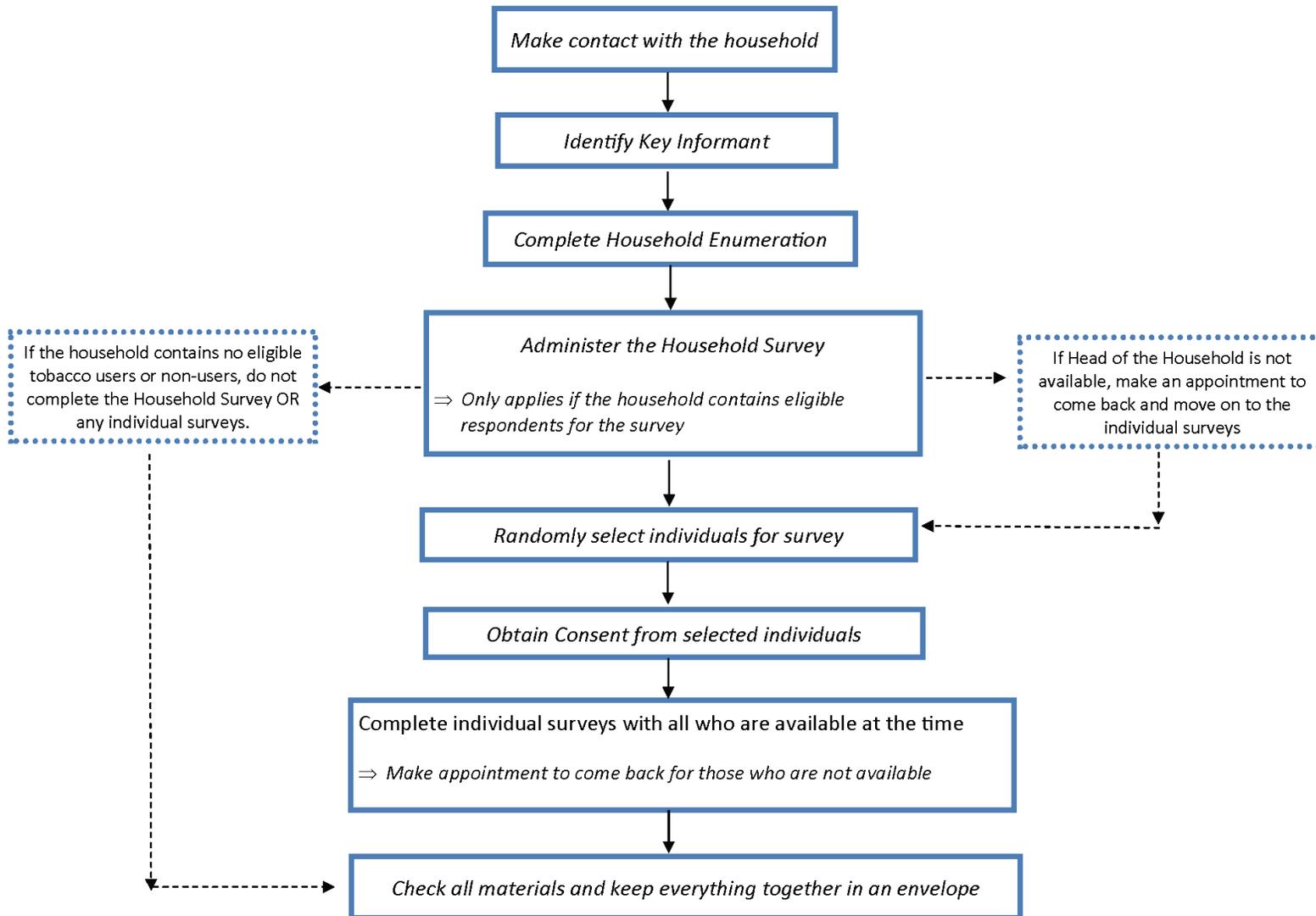


Table F1: Steps to Follow Before Approaching Households

1. Ensure permission has been obtained to survey in the area	
2. Ensure you have all materials on hand before going into the field to enumerate	<ul style="list-style-type: none">• Identification card• Household Enumeration Forms (HEFs) (One household = One HEF)• List of dwelling units to visit (prepared by your supervisor)• Die or a table of random numbers (to select members of the household to interview)• Ink pad for thumb prints in consent forms• Envelopes or folders for confidentiality• Calendar for appointment purposes• Notebook to record any problems• Map of the enumeration/survey areas if available• Blue pens, and other necessary supplies
3. Plan a route before leaving for fieldwork	
4. Ensure you approach the correct dwelling unit in the order of the dwelling unit random table	

5. Prepare the Household Enumeration Form

- Fill in ID number at the top right of Form A.
- Fill in unique household ID: Province (2 digits), District (2 digits), Division (2 digits), sub-Location (2 digits), E.A. No. (3 digits), E.A. Status (1 digit), Structure No. (5 digits), and Household No (3 digits). No other characters and symbols are allowed.
- Fill in your name.
- Fill in date of completion and sign only after you have completed the household enumeration.
- Fill in supervisor's name. They will sign and add the date after they have reviewed the form for completeness and accuracy.

Table F2: Eight Steps to Follow When Conducting Individual Interviews

Instructions: Once you have completed the Household Enumeration Form, you should thank the respondent, and continue with the following steps

<p>1. Have respondent read information letter and sign consent form</p>	<ul style="list-style-type: none"> • Respondent to read information letter (if can't read it, explain the content to respondent) • Have respondent sign two (2) copies of the consent form - one to give to respondent, one to attach to completed screener and survey questionnaire
<p>2. Use screener to determine which individual survey to use</p>	
<p>3. Clip completed screener to the individual survey questionnaire</p>	<ul style="list-style-type: none"> • Transfer respondent's ID number onto selected individual survey • Clip survey to completed screener
<p>4. Complete the survey</p>	<ul style="list-style-type: none"> • If respondent refuses to answer: • Ask if there is a more convenient date/time to return. If respondent refuses again, thank respondent for his/her time, leave and move to the next selected respondent • If respondent wishes to arrange alternate time: • Reschedule the interview and return on the agreed date and time • If respondent is not home, move on to interview another selected member of the household or if none, do NOT interview another member of the family in the household as a replacement • If interviewer cannot make interview at appointed time: • Inform the respondent and try to make another appointment
<p>5. Thank respondent and provide token of appreciation</p>	<ul style="list-style-type: none"> • Complete interview and thank respondent • Provide the appropriate token of appreciation • Let respondent know that there may be follow-up visits in subsequent years • Record the Individual Outcome Code on the HEF • Code 7 - Completed Interview
<p>6. Repeat steps 1-5 for each selected household member</p>	

7. Sign and check forms	<ul style="list-style-type: none"> • Sign the HEF and record date of completion • Check household enumeration form, consent forms, screeners, and individual survey questionnaires • Ensure all forms are completed and in order and no materials were missed or skipped • Prepare to move onto next selected household
8. Put ALL completed HEF, household surveys, consent forms, screeners, and individual surveys into one envelope/folder	<ul style="list-style-type: none"> • 1 envelope/folder = 1 household

Appendix G: Sampling Design and Weight Construction

(Please refer to the next page)

Sampling Design and Weight Construction for the International Tobacco Control (ITC) Kenya Survey

C. Boudreau^{1,2} and Y. Li^{2,3}

May 7, 2015

This technical report details the sampling design and weight construction for wave 1 of the International Tobacco Control (ITC) Kenya Survey. The ITC Kenya Survey was designed as a prospective longitudinal survey of a national representative random sample of approximately 1500 smokers and 600 non-smokers. Unfortunately, at the time this manuscript was written, the funding situation is making the future of wave 2 uncertain.

This technical report is organized as follows: section 1 describes the sampling design of the ITC Kenya Survey, section 2 details the construction of the sampling weights, and section 3 contains smoking prevalence estimates and a description of how they were computed.

1 Sampling design

The ITC Kenya Survey was designed as a prospective longitudinal study, and its sampling design was chosen to yield representative random samples of tobacco users and non-users 15 years of age and older. Respondents were interviewed in October–December 2012. All interviews were conducted face-to-face.

To qualify for the study, respondents must be 15 years old or more. Those that currently smoked cigarettes or bidis, or used any form of smokeless tobacco at least once a month were considered to be tobacco users. Those that were randomly selected completed the Smoked Tobacco Survey, the Smokeless Tobacco User Survey or Mixed User Survey depending on their tobacco consumption habit. All other qualifying respondents were considered to be non-users, and those that were randomly selected completed the Tobacco Non-User Survey.

¹Dept. of Statistics & Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada.

²Data Management Core (DMC) – ITC Project, University of Waterloo.

³Dept. of Psychology, University of Waterloo.

1 SAMPLING DESIGN

1.1 Special considerations

1. Because it is neighbouring Somalia, there are various safety concerns and other difficulties in interviewing respondents in the province of North Eastern. Instead of randomly sampling from the 4 districts in that province, it was decided to sample the district of Garissa with probability one (i.e., purposive sampling); see Patton (1990) for more information on purposive sampling. Note that Garissa is one of the two largest districts of that province.
2. The districts of Migori, Kuria West and Kuria East in the province of Nyanza are Kenya's main tobacco-growing regions. It was thus decided to sample one of those three districts. To this end, the province of Nyanza was divided into two strata: stratum NY2 consisted of the districts of Migori, Kuria West and Kuria East, whereas stratum NY1 consisted of the 18 districts of Nyanza; see table 1.
3. The districts of Mombasa, Eldoret West, Eldoret East, Wareng, Busia, Teso North and Teso South are of particular interest because of potential tobacco smuggling. It was thus decided to sample 3 of those 4 districts, one per province. The district of Mombasa is the only such district in the Coast province, and it was thus sampled with probability one. As was done for the province of Nyanza, the provinces Western and Rift Valley were each divided into two strata; see table 1.

Stratum				
#	Abbr.	Province	Districts	n_h^I
1	CE	Central	all	2
2	CO1	Coast	all except Mombasa	1
3	CO2	Coast	Mombasa	1
4	EA	Eastern	all	3
5	NA	Nairobi	all	2
6	NE	North Eastern	Garissa	1
7	NY1	Nyanza	all except those in NY2	2
8	NY2	Nyanza	Migori, Kuria West, Kuria East	1
9	RV1	Rift Valley	all except those in RV2	4
10	RV2	Rift Valley	Eldoret West, Eldoret East, Wareng	1
11	WE1	Western	all except those in WE2	2
12	WE2	Western	Busia, Teso North, Teso South	1

n_h^I = number of districts sampled in stratum h

Table 1: Strata of the ITC Kenya Survey.

1.2 Sampling frame

The sampling frame for wave 1 of the ITC Kenya Survey is the [2009 Kenya Population & Housing Census \(KPHC\)](#) conducted by the [Kenya National Bureau of Statistics \(KNBS\)](#). According to that frame, the population of Kenya was first divided into 8 provinces; see [figure 1](#). Each province was then divided into districts (or *wilaya*), with a grand total of 158 districts. Each district was then further divided into divisions (or *taarafa*), which are in turn divided into locations (or *mtaa*) and then sub-locations (or *mtaa mdogo*). Finally, each of the over 7,000 sub-locations was divided into enumeration areas (EAs). These EAs consists on average of about 100 households, but this varies quite a lot.

Our original sampling design consisted in stratifying the population, and then sampling districts (stage I), locations (stage II), sub-locations (stage III), EAs (stage IV), households (stage V) and finally individual respondents (stage VI). However, several locations contain few sub-locations; yielding selection probabilities close to 1 at stage III. For other locations, stage III selection probabilities would have been much smaller, and such a scenario would have resulted in sampling weights that would have been highly variable; thus decreasing precision. To avoid this, locations that contained too few sub-locations were pooled together or pooled with larger locations. This pooling yielded what we have called super-locations. In other words, super-locations are an artificial level between divisions and locations that we have created. The same issue arose with sub-locations containing too few EAs, and a few sub-locations were thus pooled together.

1.3 Sampling of units

To accommodate the special considerations listed in [section 1.1](#), Kenya was first divided into 12 strata; see [table 1](#). The wave 1 sample of 1500 tobacco users and 600 non-users was divided equally between 150 EAs/clusters; hence, 10 tobacco users and 4 non-users are to be randomly sampled from each EA. The sampling of units for wave 1 of ITC Kenya then proceeded as follows:

Stage I: Within stratum h ($h = 1, \dots, 12$), n_h^I districts were randomly sampled with probability proportional to population size (PPS); see [table 1](#). This stratified PPS sampling without replacement was done using the Hanurav-Vijayan algorithm; see Hanurav (1967) and Vijayan (1968) for details.

Stage II: Within selected district d ($d = 1, \dots, 21$), n_d^{II} super-locations were randomly selected using PPS sampling without replacement. In the majority of districts, 2 super-locations were sampled. For a handful of districts however, a single super-location was selected. Hence, a grand total of 37 super-locations were sampled.

1 SAMPLING DESIGN

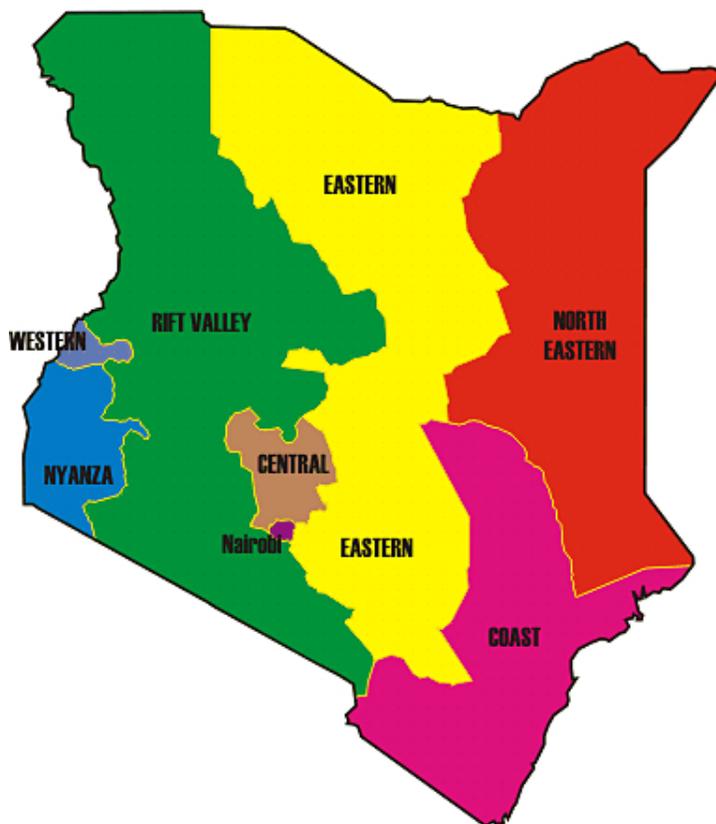


Figure 1: Provinces of Kenya.

- Stage III: 2 sub-locations were then randomly sampled with PPS in 36 of the 37 super-locations selected at stage II; whereas, 3 sub-locations were sampled via PPS in the remaining super-location (this exception was the result that sub-locations are small in that particular super-location). Since a handful of super-locations contained few sub-locations, Murthy's (1957) PPS method was used at this stage.
- Stage IV: Within selected sub-location ℓ ($\ell = 1, \dots, 75$), n_ℓ^{IV} EAs were randomly sampled with PPS. For the majority of sub-locations $n_\ell^{IV} = 2$, but there were a few instances where the EAs of a sub-location were small and thus 3 were selected instead.
- Stage V: Every households within each sampled EA was assigned a unique ID number. Those ID numbers were then randomly ordered within EA, and households were visited by interviewers following the random ordering.
- Stage VI: Finally, up to 4 tobacco users and 1 non-user per household were randomly selected and interviewed. This was repeated until both quotas were meet for



Figure 2: Districts of Kenya.

that EA; see ITC Kenya Survey Fieldwork Manual for details on stages V and VI.

A total of 5454 households were enumerated at wave 1 of the ITC Kenya Survey. Individual interviews were conducted in 1776 of those, yielding a sample of 1427 tobacco users and of 571 non-users; for a total of 1998 individual respondents.

2 Weight construction

2.1 General comments about weight construction

As with most survey weights, the ITC Kenya weights are constructed to correct and adjust for sample mis-representation caused by unequal sampling probabilities, frame error (i.e.,

2 WEIGHT CONSTRUCTION

under-coverage and multiplicity), and non-response as well as improving precision of estimates through the use of auxiliary information (e.g., tobacco usage prevalences). We briefly describe these key concepts of weight construction in this section, but refer the reader to Levy & Lemeshow (2008), chapter 16, for more detailed information.

At their base, sampling weights are defined as the inverse of selection probabilities, and thus adjust for sample misrepresentation caused by unequal sampling probabilities. For example, a tobacco user residing alone has a probability of selection twice that of a tobacco user residing with another tobacco user.

Great efforts are made to create a complete/perfect sampling frame (i.e., a frame that would include all members of the target population, without duplicate and without any erroneous inclusions¹). However, this is rarely achieved and, consequently, some members of the target population are not part of the sampling frame (i.e., have a 0 probability of being selected). This is referred to as frame under-coverage, and can result in non-coverage bias. To reduce non-coverage bias in the ITC Kenya Survey, post-stratification adjustments were performed on the sampling weights to ensure that, for each sex/age/province group for males tobacco users and sex/age group for females tobacco users, the totals of the sampling weights equal estimated number of such tobacco users based on the enumeration data; see step 3 in section 2.3. A similar adjustment was used for respondents that do not use any tobacco products; see step 4 in section 2.4.

If non-respondents behave differently than respondents, inference based solely on the sample of respondents will be biased unless adjustments are made. The greater the expected proportion of non-response, the greater this bias is likely to be. In the ITC Kenya Survey, the post-stratification adjustments described in the above paragraph also adjust for non-coverage bias. It should be noted that if data are missing completely at random (MCAR, see Little & Rubin (2002)) within each sex/age group, then non-response bias will be completely eliminated. Realistically, non-response bias is greatly reduced, but not eliminated in the ITC Kenya Survey.

All weights for the ITC Kenya Survey were computed using the statistical software SAS (<http://www.sas.com>).

2.2 Household weights

Enumeration household weights (EHWT) were computed for the 5454 households enumerated (i.e., contacted and listed) at wave 1. Computation of the EHWT weights proceeded as follows:

¹Erroneous inclusions refers to units that are not part of the target population, but included in the sampling frame.

Step 1: Each enumerated household was first assigned an enumeration area (EA) level weight $w_j^{(1)}$. This weight corresponds to the inverse of the selection probability of the j^{th} household given that its EA was sampled. Formally,

$$w_j^{(1)} = \frac{M_{k(j)}}{m_{k(j)}}$$

where j stands for the j^{th} household, $k(j)$ denotes the EA to which household j belongs, $M_{k(j)}$ is the total number of households in that EA (obtained from the 2009 census), and $m_{k(j)}$ is the number of households enumerated in that same EA. Hence, in any given EA, all households will have the same $w_j^{(1)}$ weight.

Note: For one of the 105 EAs, $m_{k(j)} > M_{k(j)}$, and the above formula could not be used as it would have resulted in $w_j^{(1)} < 1$. The computation of the $w_j^{(2)}$ weights for the households of that of EA is described in step 2 below.

Step 2: The $w_j^{(1)}$ weights were then multiplied by a factor to obtain the $w_j^{(2)}$ weights. This factor corresponds to the inverse of the selection probability of the k^{th} EA. This yielded the $w_j^{(2)}$ weights, which are formally given by

$$\begin{aligned} w_j^{(2)} &= \underbrace{\frac{N_h^I}{n_h^I N_d^{II}}}_{\text{selecting district } d} \times \underbrace{\frac{N_d^{II}}{n_d^{II} N_e^{III}}}_{\text{selecting super-location } e} \times \underbrace{\frac{N_e^{III}}{n_e^{III} N_\ell^{IV}}}_{\text{selecting sub-location } \ell} \times \underbrace{\frac{N_\ell^{IV}}{n_\ell^{IV} N_k^V}}_{\text{selecting EA } k} \times w_j^{(1)} \\ &= \frac{N_h^I}{n_h^I n_d^{II} n_e^{III} n_\ell^{IV} N_k^V} \times w_j^{(1)} \end{aligned}$$

where

- n_h^I = # of districts sampled in stratum h
- n_d^{II} = # of super-locations sampled in district d
- n_e^{III} = # of sub-locations sampled in super-location e
- n_ℓ^{IV} = # of EAs sampled in sub-location ℓ
- N_h^I = population of stratum h
- N_d^{II} = population of district d
- N_e^{III} = population of super-location e
- N_ℓ^{IV} = population of sub-location ℓ
- N_k^V = population of EA k

Notes:

- N_h^I , N_d^{II} , N_e^{III} , N_ℓ^{IV} and N_k^V were all obtained from the [2009 Kenya Population and Housing Census](#).

2 WEIGHT CONSTRUCTION

- $N_h^I / (n_h^I N_d^{II})$ is the inverse of the selection probability of district d , with population N_d^{II} , given that n_h^I districts were sampled in stratum h .
- $N_d^{II} / (n_d^{II} N_e^{III})$ is the inverse of the selection probability of super-location e , with population N_e^{III} , given that n_d^{II} super-locations were sampled in district d .
- $N_e^{III} / (n_e^{III} N_\ell^{IV})$ is the inverse of the selection probability of sub-location ℓ , with population N_ℓ^{IV} , given that n_e^{III} sub-locations were sampled in super-location e .
- $N_\ell^{IV} / (n_\ell^{IV} N_k^V)$ is the inverse of the selection probability of EA k , with population N_k^V , given that n_ℓ^{IV} EAs were sampled in sub-location ℓ .

For the EA with $m_{k(j)} > M_{k(j)}$ mentioned above, computation of the $w_j^{(2)}$ weights proceeded as follows:

- Step A) Compute $r_k = M_k / (m_k N_k^V)$ for all other EAs in the same super-location
 Step B) Compute the average of the r_k 's from step A, and let w_k^* be that average
 Step C) The $w_j^{(2)}$ weight for all households in EA k was then taken to be

$$w_j^{(2)} = \frac{N_h^I}{n_h^I n_d^{II} n_e^{III} n_\ell^{IV}} \times w_k^*$$

Step 3: Let H_j^{15} be the total number of enumerated individuals, aged 15 and older, in the j^{th} household. Multiplying H_j^{15} by $w_j^{(2)}$ and summing over all enumerated households in stratum h yields:

$$\widehat{N}_h^{15} = \sum_{j \in U_h} w_j^{(2)} H_j^{15}$$

which is an estimator of the total population, aged 15 and older, of stratum h ($h = 1, \dots, 12$).

The final step in the computation of the EHWT weights thus consists in multiplying the $w_j^{(2)}$ weights by a factor to ensure that

$$\sum_{j \in U_h} w_j^{\text{EHWT}} H_j^{15} = N_h^{15}$$

where N_h^{15} is the total population, aged 15 and older, of stratum h obtained from the 2009 census; see table 2. Consequently,

$$w_j^{\text{EHWT}} = w_j^{(2)} N_h^{15} / \widehat{N}_h^{15}$$

Stratum	Population (N_h^{15})
NA	2,187,366
CE	2,805,639
CO2	352,605
EA	3,300,875
NE	103,426
NY2	172,634
RV2	523,494
WE2	302,837
CO1	1,564,694
NY1	2,770,572
RV1	4,952,439
WE1	1,988,495

Table 2: Population, aged 15 and older, by stratum.

Interviewed household weights (IHWT) were computed for the 1776 households (i.e., 32.6% of the households enumerated at wave 1) where one or more (i.e., up to 4) respondents completed one of the individual surveys (i.e., the cigarette, smokeless, mixed tobacco or non-user survey). To this end, we first need to divide the households into 2 groups: user households and non-user households. A household is said to be a non-user household if none of its occupants, aged 15 and older, use tobacco products; otherwise, it is classified as a user household. Computation of the IHWT weights proceeded as follows,

$$w_j^{\text{IHWT}} = \begin{cases} w_j^{\text{EHWT}} \times \frac{UH_{k(j)}^*}{UH_{k(j)}} & \text{if household } j \text{ is a tobacco user household} \\ w_j^{\text{EHWT}} \times \frac{NH_{k(j)}^*}{NH_{k(j)}} & \text{if household } j \text{ is a non-user household} \end{cases}$$

where w_j^{EHWT} is the EHWT weight of household j computed above, $UH_{k(j)}$ is the number of user households interviewed in the EA of household j , and $UH_{k(j)}^*$ is the number of user households enumerated in that same EA. Similarly, for non-user households, $NH_{k(j)}$ is the number of non-user households interviewed in the EA of household j , and $NH_{k(j)}^*$ is the number of non-user households enumerated in that same EA.

2.3 Individual weights for tobacco users

Weights were computed for the 1427 respondents that completed one of the following 3 individual surveys: Smoked Tobacco Survey, Smokeless Tobacco User Survey or Mixed User Survey. Computation of those weights proceeded as follows:

2 WEIGHT CONSTRUCTION

Step 1: Each tobacco user was first assigned a within-household weight $w_i^{(1)}$, which corresponds to the inverse of his/her selection probability amongst eligible tobacco users of the same sex in his/her household. Formally, the $w_i^{(1)}$ weights are given by

$$w_i^{(1)} = \begin{cases} \frac{\#\text{male users}_{j(i)}}{\#\text{male users intvw}_{j(i)}} & \text{if } i^{\text{th}} \text{ respondent is a male} \\ \frac{\#\text{female users}_{j(i)}}{\#\text{female users intvw}_{j(i)}} & \text{if } i^{\text{th}} \text{ respondent is a female} \end{cases}$$

where $j(i)$ denotes the household in which respondent i lives, $\#\text{male users}_{j(i)}$ is the number of eligible male tobacco users in that same household, $\#\text{male user intvws}_{j(i)}$ is the number of such household members that were interviewed, and similarly for female tobacco users. Note: $w_i^{(1)}$ was equal to 1 for most respondents, and can only be greater than 1 in cases where more than 4 eligible tobacco users resided in the same household.

Step 2: Each $w_i^{(1)}$ weight was then multiplied by the corresponding IHWT weight. The resulting weights are labelled $w_i^{(2)}$, and are formally defined as

$$w_i^{(2)} = w_i^{(1)} \times w_{j(i)}^{\text{IHWT}}$$

where $w_{j(i)}^{\text{IHWT}}$ is the IHWT weight of the j^{th} household (computed in section 2.2).

Step 3: A post-stratification adjustment was then performed to calibrate the $w_i^{(2)}$ weights to estimated prevalences of tobacco usage by sex/age/province groups; see section 3 for a description of how those prevalences were estimated. To this end, age was divided into 4 intervals; i.e., $[15, 25)$, $[25, 35)$, $[35, 50)$ and $[50, 100)$. Males were further subdivided according to the province they live in. This yielded the first 28 sex/age/province cells of table 3, where some provinces and/or age groups were collapsed because they contained too few respondents. Females were not subdivided by provinces, as the prevalence of tobacco usage is very low amongst women. This yielded the remaining 4 sex/age group cells of table 3.

For respondents in cell C_k , this post-stratification adjustment consisted in multiplying their $w_i^{(2)}$ weights by a factor $\widehat{N}_k^{\text{user}}/t_k$ to produce calibrated $w_i^{(3)}$ weights. These $w_i^{(3)}$ weights are such that their sum over all respondents in cell C_k is equal to $\widehat{N}_k^{\text{user}}$, the estimated number of tobacco users, aged 15 and older, in that cell. Formally,

$$w_i^{(3)} = w_i^{(2)} \times \frac{\widehat{N}_k^{\text{user}}}{t_k} = w_i^{(2)} \times \frac{\widehat{N}_k^{\text{user}}}{\sum_{i \in C_k} w_i^{(2)}}$$

where $\widehat{N}_1^{\text{user}}, \dots, \widehat{N}_{32}^{\text{user}}$ are given in column 6 of table 3, and C_k is the set of all respondents in cell k ($k = 1, \dots, 32$).

Step 4: Finally, the $w_i^{(3)}$ weights were rescaled for use in pooled regression or logistic regression analyses where age, gender and tobacco use status are covariates, as well as to facilitate comparisons with other ITC countries. To this end, the 1427 respondents were divided into 2 groups: male tobacco users ($g = 1$), and female tobacco users ($g = 2$). The weights were then rescaled to have mean equal to 1 in each group. This yielded the $w_i^{(4)}$ weights, which are formally defined as

$$w_i^{(4)} = w_i^{(3)} \times \frac{n_g}{\sum_{i \in S_g} w_i^{(3)}},$$

where S_g is the set of all respondents in group g , and n_g is the size of that sample; i.e.,

$$n_g = \begin{cases} 1183 & \text{if } g = 1 \\ 244 & \text{if } g = 2 \end{cases}$$

Note: the coefficient of variation (cv) of the $w_i^{(4)}$ weights is 1.42 for male tobacco users ($g = 1$), and 0.78 for female tobacco users ($g = 2$).

2.4 Individual weights for non-users

Weights were computed for the the 571 respondents that do not use any tobacco product, and thus completed the individual Tobacco Non-User Survey. Computation of those weights proceeded alike that of the weights for tobacco users; i.e.:

Step 1: Each non-user was first assigned a within-household weight $w_i^{(1)}$, which corresponds to the inverse of his/her selection probability amongst eligible non-users in his/her household. Formally, the $w_i^{(1)}$ weights are given by

$$w_i^{(1)} = \frac{\#\text{non-user}_{j(i)}}{\#\text{non-user intvw}_{j(i)}}$$

where $j(i)$ denotes the household in which respondent i lives, $\#\text{non-user}_{j(i)}$ is the number of household members aged 15 and older that do not use any tobacco products, and $\#\text{non-user intvw}_{j(i)}$ is the number of such household members that were interviewed (i.e., completed the individual Tobacco Non-User Survey).

Notes:

- $\#\text{non-user}_{j(i)}$ was capped at 3 to control the variability of the weights.
- According to the sampling plan, $\#\text{non-user intvw}_{j(i)}$ should be 1 for all households. However, there were few instances where 2 non-users from the same household were interviewed.

2 WEIGHT CONSTRUCTION

Step 2: Each $w_i^{(1)}$ weight was then multiplied by the corresponding IHWT weight. The resulting weights are labelled $w_i^{(2)}$, and are formally defined as

$$w_i^{(2)} = w_i^{(1)} \times w_{j(i)}^{\text{IHWT}}$$

where $w_{j(i)}^{\text{IHWT}}$ is the IHWT weight of the j^{th} household (computed in section 2.2).

Step 3: Since a non-user residing with one or more tobacco users could only be selected while the non-user quota was opened, the $w_i^{(2)}$ weights of such non-users were multiplied by the estimated probability that the non-user quota was opened when interviewing a random household in EA k . This yielded the $w_i^{(3)}$ weights, which are formally defined as

$$w_i^{(3)} = \begin{cases} w_i^{(2)} & \text{if } i^{\text{th}} \text{ respondent lives in an household with} \\ & \text{only non-users} \\ w_i^{(2)} \times \frac{NH_{k(j)}^*}{NH_{k(j)}} & \text{if } i^{\text{th}} \text{ respondent lives in an household with 1} \\ & \text{or more tobacco users} \end{cases}$$

where $NH_{k(j)}$ and $NH_{k(j)}^*$ were defined in the computation of the IHWT weights.

Step 4: A post-stratification adjustment was then performed to calibrate the $w_i^{(3)}$ weights to the estimated number of individuals who do not use any tobacco products by sex/age groups; see section 3 for a description of how those numbers were estimated. To this end, age was divided into 4 intervals; i.e., [15, 25), [25, 35), [35, 50), [50, 100). This yielded the 8 sex/age cells of table 4. Note that we did not further divide by province, as our sample of non-users is much smaller than our sample of tobacco users. Hence, further subdividing by province would have resulted in cells with too few respondents, which could have in turn resulted in less stable weights.

For respondents in cell C_k , this post-stratification adjustment consisted in multiplying their $w_i^{(3)}$ weights by a factor $\widehat{N}_k^{\text{non-user}}/t_k$ to produce calibrated $w_i^{(4)}$ weights. These $w_i^{(3)}$ weights are such that their sum over all respondents in cell C_k is equal to $\widehat{N}_k^{\text{user}}$, the estimated number of individuals, aged 15 and older, that do not use any tobacco products in that cell. Formally,

$$w_i^{(4)} = w_i^{(3)} \times \frac{\widehat{N}_k^{\text{non-user}}}{t_k} = w_i^{(3)} \times \frac{\widehat{N}_k^{\text{non-user}}}{\sum_{i \in C_k} w_i^{(3)}}$$

where $\widehat{N}_k^{\text{non-user}}, \dots, \widehat{N}_k^{\text{non-user}}$ are given in column 5 of table 4, and C_k is the set of all respondents in cell k ($k = 1, \dots, 8$).

Step 5: Finally, the $w_i^{(4)}$ weights were rescaled for use in pooled regression or logistic regression analyses where age, gender and tobacco use status are covariates, as

well as to facilitate comparisons with other ITC countries. To this end, the 571 respondents were divided into 2 groups: non-users living with 1 or more tobacco users ($g = 1$), and non-users living with only other non-users ($g = 2$). The weights were then rescaled to have mean equal to 1 in each group. This yielded the $w_i^{(5)}$ weights, which are formally defined as

$$w_i^{(5)} = w_i^{(4)} \times \frac{n_g}{\sum_{i \in S_g} w_i^{(4)}}$$

where S_g is the set of all respondents in group g , and n_g is the size of that sample; i.e.,

$$n_g = \begin{cases} 103 & \text{if } g = 1 \\ 468 & \text{if } g = 2 \end{cases}$$

Note: the coefficient of variation (cv) of the $w_i^{(4)}$ weights is 1.31 for non-users living with 1 or more tobacco users ($g = 1$), and 1.47 for non-users living with only other non-users ($g = 2$).

3 Prevalence estimates

This section details how estimates for the prevalences of tobacco usage, per gender/age/ province group, and the number of tobacco users were computed. These estimates are given in table 3, and were used to calibrate the sampling weights of tobacco users (see step 3 of section 2.3). This section also describes how we estimated the numbers of individuals, per gender/age group, who do not use any tobacco product, and the corresponding prevalences. These estimates are given in table 4, and were used to calibrate the sampling weights of non-users (see step 4 of section 2.4). All estimates are based on the 5454 household enumeration forms completed at wave 1 of the ITC Kenya Survey.

Estimation of population totals and averages from survey data is usually very straightforward. In the present case however, we had the additional difficulty that age was missing for 442 individuals, and that tobacco use status was missing for another 27 individuals². Failure to take this into account would have led to under-estimation of the number of tobacco users, and likely bias prevalence estimates.

In section 3.1, we describe how calculations would have proceed if there were no missing values; whereas, in section 3.2, we describe the procedure we actually used to obtain the figures in tables 3 and 4, which take into accounts the missing status and ages.

²There was also 13 individuals that had both age and tobacco usage status missing. However, to keep the estimation procedure of section 3.2 simple, these individuals were simply excluded. Since there are so few such individuals, doing so has effectively no impact.

3 PREVALENCE ESTIMATES

3.1 Assuming no missing values

Let k be one of the sex/province(s)/age group combinations for males, or sex/age group combinations for females, of table 3. The estimated number of tobacco users for the k^{th} combination ($k = 1, \dots, 32$) is given by

$$\tilde{N}_k^{\text{user}} = \sum_{j \in U_h} w_j^{\text{EHWT}} HU_j^k$$

where h denotes the province (or provinces since some were merged) corresponding to the k^{th} combination, w_j^{EHWT} is the EHWT of the j^{th} household (computed in section 2.2), and HU_j^k is the number of enumerated tobacco users in the household j that fall within the sex/age group corresponding to the k^{th} combination. Note that $HU_j^k = 0$ if household j contains no such individuals. Lastly, the sum is over all enumerated households in the h^{th} province(s) for males, and over all enumerated households nationwide for females (since all 8 provinces were pooled together). Likewise, the estimated total number of individuals for the k^{th} combination is given by

$$\tilde{N}_k^{\text{total}} = \sum_{j \in U_h} w_j^{\text{EHWT}} H_j^k$$

where H_j^k is the total number of enumerated individuals in the j^{th} household that fall within the sex/age group corresponding to the k^{th} combination. Combining $\tilde{N}_k^{\text{user}}$ and $\tilde{N}_k^{\text{total}}$, the estimated prevalence of tobacco usage for the k^{th} combination is given by

$$\tilde{p}_k^{\text{user}} = \frac{\tilde{N}_k^{\text{user}}}{\tilde{N}_k^{\text{total}}}$$

Now, let k be one of the sex/age group combinations of table 4. Using the same logic, the estimated number of non-users for the k^{th} combination ($k = 1, \dots, 8$) is given by

$$\tilde{N}_k^{\text{non-user}} = \sum_{j \in U} w_j^{\text{EHWT}} HN_j^k$$

where HN_j^k is the number of enumerated individuals in the j^{th} household that do not use any tobacco product and fall within the sex/age group corresponding to combination k , and the sum is over all enumerated households nationwide (since all 8 provinces were pooled together for estimating non-users). Note: $H_j^k = HU_j^k + HN_j^k$. The formula for $\tilde{N}_k^{\text{total}}$ remains the same, but the the subscript k now refers to the k^{th} combination of table 4; hence, the sum is now over all enumerated households nationwide. After computing the new $\tilde{N}_k^{\text{total}}$, the estimated proportion of individuals that do not use any tobacco product for the k^{th} combination is given by

$$\tilde{p}_k^{\text{non-user}} = \frac{\tilde{N}_k^{\text{non-user}}}{\tilde{N}_k^{\text{total}}}$$

3.2 With missing values

We now incorporate the $442 + 27 = 469$ individuals with missing values. To this end, let HS_j^k be the number of enumerated individuals in the j^{th} household that fall within the sex/age group corresponding to k^{th} combination of table 3, but for whom the tobacco usage status is unknown. In addition, let

$$HA_j^g = \begin{cases} \# \text{ of enumerated males in the } j^{\text{th}} \text{ household with missing age} & \text{if } g = 1 \\ \# \text{ of enumerated females in the } j^{\text{th}} \text{ household with missing age} & \text{if } g = 2 \end{cases}$$

HA_j^g needs to be divided into those that are tobacco users and those that do not use any tobacco product. Hence, let HAU_j^g be the number of enumerated male ($g = 1$)/female ($g = 2$) tobacco users in the j^{th} household with missing age, and similarly for HAN_j^g ; thus, $HAU_j^g + HAN_j^g = HA_j^g$. Let $S_1(g, a, d)$ be the set of all enumerated male ($g = 1$)/female ($g = 2$) respondents in age group a ($a = [15, 25), [25, 35), [35, 50)$ or $[50, 100)$) residing in district d ($d = 1, \dots, 21$). Let $\theta_d^{a,g}$ be the proportion of those individuals that use tobacco products. In addition, let $S_2(g, d)$ be the set of all enumerated male ($g = 1$)/female ($g = 2$) respondents residing in district d . Let $\beta_d^{(a,g)}$ be the proportion of those individuals that are in age group a ; thus, $\sum_a \beta_d^{(a,g)} = 1$. The estimated number of tobacco users for the k^{th} combination ($k = 1, \dots, 32$) of table 3 is given by

$$\widehat{N}_k^{\text{user}} = \sum_{j \in U_h} w_j^{\text{EHWT}} \left(HU_j^k + HAU_j^g \beta_{d(j)}^{(a,g)} + HS_j^k \theta_{d(j)}^{(a,g)} \right)$$

where $d(j)$ denotes the district where the j^{th} household is located, and (a, g) are the sex/age group corresponding to the k^{th} combination. As with $\widehat{N}_k^{\text{user}}$, the sum is over all enumerated households in the h^{th} province(s) for males, and over all enumerated households nationwide for females. Similarly, the estimated total number of individuals for the same k^{th} combination is given by

$$\widehat{N}_k^{\text{total}} = \sum_{j \in U_h} w_j^{\text{EHWT}} \left(H_j^k + HA_j^g \beta_{d(j)}^{(a,g)} + HS_j^k \right)$$

Combining $\widehat{N}_k^{\text{user}}$ and $\widehat{N}_k^{\text{total}}$, the estimated prevalence of tobacco usage for the k^{th} combination of table 3 is given by

$$\hat{p}_k^{\text{user}} = \frac{\widehat{N}_k^{\text{user}}}{\widehat{N}_k^{\text{total}}}$$

Now, let k be one of the sex/age group combinations of table 4. Using the same logic, the estimated number of non-users for the k^{th} combination ($k = 1, \dots, 8$) is given by

$$\widehat{N}_k^{\text{non-user}} = \sum_{j \in U} w_j^{\text{EHWT}} \left(HN_j^k + HAN_j^g \beta_{d(j)}^{(a,g)} + HS_j^k (1 - \theta_{d(j)}^{(a,g)}) \right)$$

REFERENCES

As with \hat{p}_k^{user} , the formula for \hat{N}_k^{total} remains the same, but the subscript k now refers to the k^{th} combination of table 4; hence, the sum is now over all enumerated households nationwide. After computing the new \hat{N}_k^{total} , the estimated proportion of individuals that do not use any tobacco product for the k^{th} combination of table 4 is given by

$$\hat{p}_k^{\text{non-user}} = \frac{\hat{N}_k^{\text{non-user}}}{\hat{N}_k^{\text{total}}}$$

Acknowledgements

Core funding for the ITC Project is provided by the U.S. National Cancer Institute (NCI; grants P50 CA111236 & P01 CA138389), the Canadian Institutes of Health Research (CIHR; grants 79551 & 115016), and by the Ontario Institute for Cancer Research (OICR).

References

- Hanurav, T. V. (1967), ‘Optimum utilization of auxiliary information: π_{ps} sampling of two units from a stratum’, *Journal of the Royal Statistical Society, Series B* **29**, 374–391.
- Levy, P. S. & Lemeshow, S. (2008), *Sampling of Populations: Methods and Applications*, 4 edn, John Wiley & Sons, Hoboken, NJ.
- Little, R. J. A. & Rubin, D. B. (2002), *Statistical Analysis with Missing Data*, 2 edn, John Wiley & Sons, Hoboken, NJ.
- Murthy, M. N. (1957), ‘Ordered and unordered estimators in sampling without replacement’, *Sankhyā* **18**, 379–390.
- Patton, M. Q. (1990), *Qualitative evaluation and research methods*, 2 edn, Sage, Newbury Park, CA.
- Vijayan, K. (1968), ‘An exact π_{ps} sampling scheme: Generalization of a method of Hanurav’, *Journal of the Royal Statistical Society, Series B* **30**, 556–566.

k	sex	province(s)	age	\hat{p}_k^{user}	\hat{N}_k^{user}
1	male	Nairobi	[15-24)	13.3%	32,660
2	male	Nairobi	[25-34)	27.9%	126,564
3	male	Nairobi	[35-49)	28.9%	80,005
4	male	Nairobi	[50-100)	37.1%	27,115
5	male	Central	[25-34)	43.3%	128,117
6	male	Central	[35-49)	41.2%	147,073
7	male	Central	[50-100)	42.1%	173,347
8	male	Coast	[15-24)	35.0%	115,919
9	male	Coast	[25-34)	49.1%	197,905
10	male	Coast	[35-49)	45.1%	123,804
11	male	Eastern	[15-24)	15.6%	80,619
12	male	Eastern	[25-34)	37.8%	152,534
13	male	Eastern	[35-49)	44.2%	162,705
14	male	Eastern	[50-100)	55.6%	223,857
15	male	North Eastern	[15-24)	17.1%	3,561
16	male	North Eastern	[25-49)	26.2%	7,252
17	male	North Eastern	[50-100)	37.8%	3,324
18	male	Nyanza	[25-34)	18.0%	67,127
19	male	Nyanza	[35-49)	25.2%	74,994
20	male	Nyanza	[50-100)	27.9%	86,469
21	male	Rift Valley	[15-24)	10.2%	89,150
22	male	Rift Valley	[25-34)	28.3%	256,306
23	male	Rift Valley	[35-49)	33.7%	206,973
24	male	Rift Valley	[50-100)	40.8%	165,554
25	male	Western	[25-34)	18.9%	51,447
26	male	Western	[35-49)	26.7%	68,511
27	male	Coast+Western	[50-100)	23.8%	74,664
28	male	Central+Nyanza+Western	[15-24)	3.8%	17,473
29	female	all	[15-24)	0.9%	31,758
30	female	all	[25-34)	3.2%	89,437
31	female	all	[35-49)	4.2%	90,049
32	female	all	[50-100)	19.9%	362,613

Table 3: Estimated numbers of individuals that use one or more tobacco products, per sex/province(s)/age group for males and sex/age group for females, and corresponding prevalences.

REFERENCES

k	sex	age	$\hat{p}_k^{\text{non-user}}$	$\hat{N}_k^{\text{non-user}}$
1	male	[15-24)	88.4%	2,789,174
2	male	[25-34)	68.7%	2,140,830
3	male	[35-49)	64.7%	1,587,803
4	male	[50-100)	60.8%	1,171,712
5	female	[15-24)	99.1%	3,567,341
6	female	[25-34)	96.8%	2,691,307
7	female	[35-49)	95.8%	2,038,155
8	female	[50-100)	80.1%	1,463,146

Table 4: Estimated numbers of individuals that do not use any tobacco product, per sex/age group, and corresponding prevalences.