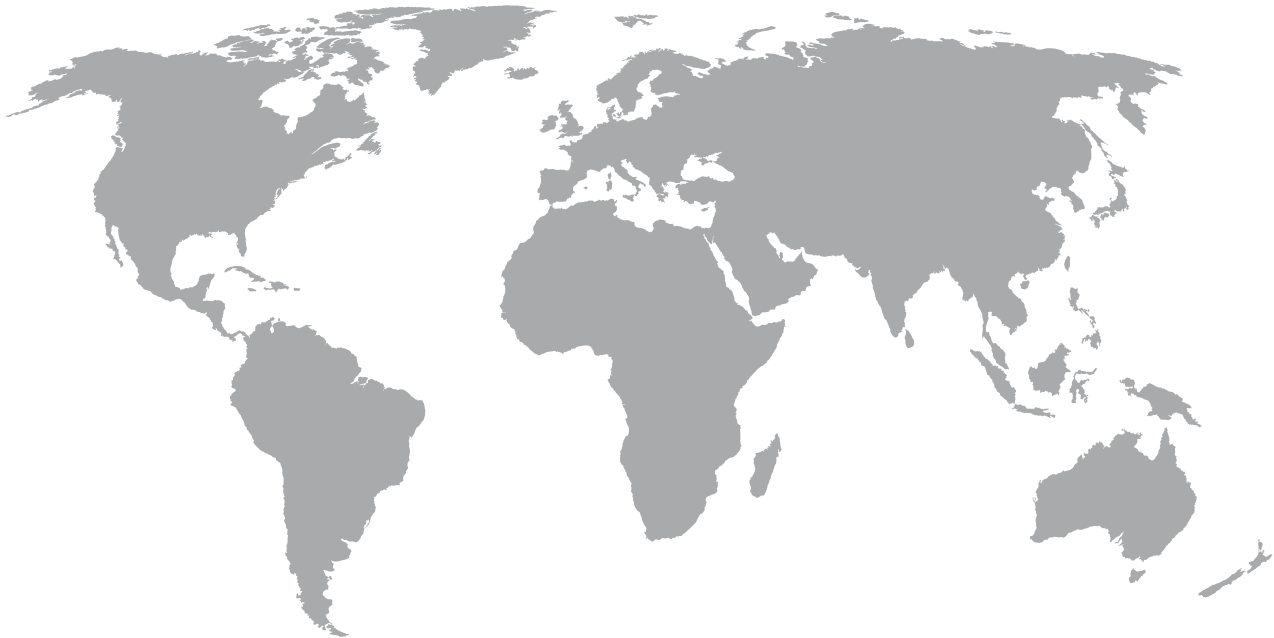




Global Adult Tobacco Survey (GATS)



Assurance qualité : directives et documentation

**Enquête mondiale sur le tabagisme
chez les adultes (GATS)
Assurance qualité :
directives et documentation**

Septembre 2020

Enquête mondiale sur le tabagisme chez les adultes (GATS)

Protocole standard complet

Questionnaire GATS

Questionnaire principal avec questions facultatives
Spécifications question par question

Plan d'échantillonnage pour l'enquête GATS

Manuel d'échantillonnage
Manuel de pondération de l'échantillon

Mise en œuvre de l'enquête GATS sur le terrain

Manuel d'enquête sur le terrain
Manuel de supervision sur le terrain
Manuel de cartographie et de constitution des listes

Gestion des données de l'enquête GATS

Guide de programmation sur General Survey System
Spécifications pour la programmation du questionnaire principal
Plan de mise en œuvre pour la gestion des données
Guide de formation à la gestion des données

Assurance qualité de l'enquête GATS : directives et documentation

Outils d'analyse et de rédaction du rapport de l'enquête GATS

Modèles de fiches d'information
Rapport national : plan et directives de mise en tableau
Définitions des indicateurs

Publication et diffusion des données de l'enquête GATS

Politique de publication des données
Diffusion des données : conseils relatifs à la publication initiale des données

Pour citer ce document :

Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Assurance qualité : directives et documentation*. Atlanta, Géorgie (États-Unis d'Amérique d'Amérique) : Centers for Disease Control and Prevention, 2020.

Remerciements

Organisations partenaires de l'enquête GATS

- Centers for Disease Control and Prevention (CDC, États-Unis d'Amérique)
- CDC Foundation
- Institut de santé publique Johns Hopkins Bloomberg (JHSPH)
- RTI International
- Organisation mondiale de la Santé (OMS)

Soutien financier

Un soutien financier a été apporté par la *Bloomberg Initiative to Reduce Tobacco Use* (Initiative Bloomberg de lutte contre le tabagisme), un programme de Bloomberg Philanthropies, par l'intermédiaire de la CDC Foundation.

Attention : les opinions exprimées dans le présent manuel ne reflètent pas nécessairement celles des organisations partenaires de l'enquête GATS.

Table des matières

Chapitre	Page
1. Introduction	1-1
1.1 Présentation générale de l'Enquête mondiale sur le tabagisme chez les adultes.....	1-2
1.2 Utilisation du présent manuel	1-2
1.3 Assurance qualité de l'enquête GATS	1-3
2. Plan du processus d'assurance qualité.....	2-1
3. Assurance qualité : Actions à mener avant la collecte des données.....	3-1
3.1 Conception du questionnaire et processus d'examen	3-1
3.2 Programmation et contrôle de versions	3-2
3.3 Test préliminaire.....	3-5
3.4 Finalisation du questionnaire	3-6
3.5 Transfert et gestion des données.....	3-7
3.6 Plan d'échantillonnage et processus d'examen.....	3-7
3.7 Préparation du fichier de sélection de l'échantillon-maître	3-7
3.8 Ateliers de formation	3-8
4. Assurance qualité : Collecte et gestion des données.....	4-1
4.1 Collecte de données sur le terrain : Configuration et maintenance	4-1
4.2 Gestion des données du Centre national des données	4-8
5. Assurance qualité : Actions à mener après la collecte des données.....	5-1
5.1 Nettoyage et préparation des données pour le calcul des pondérations de l'échantillon.....	5-2
5.2 Mesures de la qualité : échantillonnage, erreur d'échantillonnage et pondération.....	5-6
5.3 Mesures de la qualité : couverture, non-réponse et autres erreurs non liées à l'échantillonnage	5-12
5.4 Examen formel de la qualité statistique	5-18
5.5 Création du fichier de données analytiques	5-20
5.6 Création du fichier de données destiné à une utilisation publique.....	5-20
6. Critères en matière de données et de documentation	6-1
7. Bibliographie	7-1
Annexe A : Glossaire	A-1
Annexe B : Actions à mener avant la collecte des données	B-1
B.1 Adaptation du questionnaire GATS et processus d'examen	B-1
B.2 Processus de programmation du questionnaire de l'enquête GATS.....	B-2
B.3 Processus de configuration du programme GSS sur les dispositifs portables	B-2
Annexe C : Actions à mener après la collecte des données	C-1

C.1	Codes de décision finale et calcul des taux de réponse.....	C-1
C.2	Tendance des ajustements d'étalonnage des pondérations par stratification a posteriori au sein des cellules d'ajustement	C-5
C.3	Effet de la variable de pondération de l'échantillon sur la précision des estimations de l'enquête.....	C-7
C.4	Effet du plan de sondage global sur la précision des estimations de l'enquête et sur l'homogénéité intragroupe dans les unités primaires d'échantillonnage (UPE) des principales estimations correspondantes de l'enquête	C-8
C.5	Marge d'erreur pour les principales estimations de l'enquête	C-10
C.6	Estimation des erreurs d'échantillonnage	C-13
C.7	Taux de couverture de la base de ménages.....	C-21

Documents

Numéro	Page
Document 2-1. Diagramme des flux relatifs à l'assurance qualité.....	2-1

1. Introduction

Le tabagisme est l'une des principales causes évitables de maladies et de mortalité précoce à travers le monde, qui touche environ 1,4 milliard de personnes âgées de 15 ans ou plus¹. En outre, plus de 8 millions de personnes meurent chaque année de maladies liées au tabac². Si les tendances actuelles se poursuivent, le tabagisme pourrait tuer un milliard de personnes avant la fin du siècle, et l'on estime que plus de trois quarts de ces décès surviendront dans des pays à revenu faible et intermédiaire³. Un dispositif de surveillance systématique et efficace est essentiel pour assurer le suivi et la gestion de cette épidémie.

L'**Enquête mondiale sur le tabagisme chez les adultes** (Global Adult Tobacco Survey, GATS), réalisée dans le cadre du système mondial de surveillance du tabagisme (Global Tobacco Surveillance System, GTSS), constitue une référence mondiale en matière de surveillance systématique de la consommation de tabac chez les adultes et de suivi des principaux indicateurs de la lutte antitabac. Il s'agit d'une enquête représentative à l'échelle nationale, menée auprès des ménages composés de personnes de 15 ans ou plus, suivant un questionnaire principal standardisé et un plan d'échantillonnage, de collecte et de gestion des données examinées et approuvées par des experts internationaux. Elle vise à renforcer la capacité des pays à concevoir, mettre en œuvre et évaluer leurs programmes de lutte antitabac.

Afin de maximiser l'efficacité des données recueillies dans le cadre de l'enquête GATS, une série de manuels a été créée. Ces manuels ont pour objectif de présenter aux pays les exigences de base et d'offrir des recommandations relatives à l'élaboration et au déroulement de l'enquête à chaque étape du processus GATS. Ils apportent également des conseils sur la façon dont un pays donné peut adapter certains éléments du protocole GATS afin de maximiser l'utilité des données à l'échelle nationale. Il est vivement recommandé de suivre le protocole standard pour assurer la cohérence et la comparabilité des données entre les pays.

Les manuels GATS fournissent des conseils systématiques sur l'élaboration et le déroulement de l'enquête.

¹ Organisation mondiale de la Santé (OMS), WHO report on the global tobacco epidemic, 2019: Offer help to quit tobacco use. Genève, Suisse : Organisation mondiale de la Santé, Genève (Suisse), 2019. Disponible à l'adresse suivante : <https://apps.who.int/iris/bitstream/handle/10665/326043/9789241516204-eng.pdf?ua=1>.

² GBD 2017 Risk Factor Collaborators, « Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017 ». Institute for Health Metrics and Evaluation, Seattle, Washington (États-Unis d'Amérique), 2018.

³ Mathers, C. D. et Loncar, D., « Projections of Global Mortality and Burden of Disease from 2002 to 2030 ». *PLoS Medicine*, vol. 3, n° 11, e442, 2006.

1.1 Présentation générale de l'Enquête mondiale sur le tabagisme chez les adultes

L'enquête GATS vise à produire des estimations nationales et infranationales concernant les adultes dans différents pays. Elle cible l'ensemble des hommes et des femmes de 15 ans ou plus non placés en institution, et qui considèrent le pays concerné comme leur lieu de résidence habituel. Tous les membres de la population cible seront échantillonnés dans les ménages qu'ils considèrent comme leur lieu de résidence habituel.

L'enquête GATS repose sur une méthode d'échantillonnage par zones géographiques en plusieurs étapes permettant d'identifier les ménages qui seront contactés par les enquêteurs de terrain. Un pays est d'abord divisé en unités primaires d'échantillonnage, puis en segments, puis en ménages. Un échantillon de ménages est ensuite choisi au hasard pour participer à l'enquête GATS.

Les entretiens de l'enquête GATS se décomposent en deux parties : un *questionnaire destiné aux ménages* et un *questionnaire individuel*. Le *questionnaire destiné aux ménages* (sélection des ménages) et le *questionnaire individuel* (entretien individuel) seront administrés à l'aide d'un dispositif électronique de collecte de données.

Les entretiens de l'enquête GATS se décomposent en deux parties : un questionnaire destiné aux ménages et un questionnaire individuel. Ces questionnaires sont administrés à l'aide d'un dispositif électronique de collecte de données.

À chaque adresse comprise dans l'échantillon, les enquêteurs de terrain demanderont à un adulte faisant partie du ménage de répondre au *questionnaire destiné aux ménages*. L'objectif de ce questionnaire est de déterminer si le ménage sélectionné répond aux critères d'admissibilité de l'enquête GATS et de dresser la liste de tous les membres admissibles au sein du ménage.

Une fois la liste terminée, l'une des personnes sera choisie au hasard pour répondre au *questionnaire individuel*. Le *questionnaire individuel* porte sur le profil sociodémographique, le tabagisme, les cigarettes électroniques, le tabac sans fumée, le sevrage, le tabagisme passif, l'aspect économique, les médias, mais aussi sur les connaissances, les mentalités et les perceptions au sujet du tabac.

1.2 Utilisation du présent manuel

Le présent manuel fournit des orientations concernant l'évaluation et l'assurance qualité des données collectées dans le cadre de l'enquête GATS.

Il se présente comme un complément aux manuels GATS existants, en ce sens qu'il ne comportera pas de descriptions des procédures documentées ailleurs. Au lieu de cela, il décrit les étapes du processus d'assurance qualité qui doivent avoir lieu tout au long de la mise en œuvre de l'enquête GATS.

L'application de ces directives d'assurance qualité est une condition essentielle à la réussite de cette enquête. Ce manuel est destiné à être utilisé par toutes les personnes impliquées dans la mise en œuvre et dans le processus d'assurance qualité de l'enquête GATS, y compris le personnel informatique et les statisticiens responsables de l'échantillonnage et de la pondération. Un glossaire des termes utilisés est fourni en *annexe A*.

Dans ce document, le processus d'assurance qualité de l'enquête GATS est segmenté en trois étapes successives :

les actions à mener avant la collecte des données, la collecte et la gestion des données, et les actions à mener après la collecte de données. Les mesures d'assurance qualité varient au fil des différentes

étapes du processus d'assurance qualité, et des directives sont fournies pour chaque étape. Le *Chapitre 2* présente en préambule un plan du processus, décrivant la portée globale des activités d'assurance qualité.

Le *Chapitre 3 (Actions à mener avant la collecte des données)* inclut des sections décrivant les mesures d'assurance qualité qu'il convient d'appliquer au moment de la conception et de l'examen du questionnaire adapté au contexte national, de la programmation du questionnaire pour les dispositifs portables utilisés durant la collecte de données, du plan d'échantillonnage, et de la création du fichier de sélection de l'échantillon-maître.

Le *Chapitre 4 (Collecte et gestion des données)* est divisé en trois parties. La première traite des activités d'assurance qualité réalisées sur le terrain, telles que la mise en place et le suivi du système de gestion des enquêtes et du questionnaire. Elle inclut également des sous-parties consacrées à la gestion des données et définit les normes à respecter en matière d'établissement de rapports réguliers durant les activités de terrain. La deuxième partie décrit les opérations de gestion des données et d'assurance qualité qui s'appliquent une fois les entretiens réalisés et les données chargées sur le site central. La troisième partie traite des enjeux d'agrégation et de contrôle qualité qui interviennent au niveau national. Elle décrit également la préparation du fichier de données brutes pour le nettoyage.

Le *Chapitre 5 (Actions à mener après la collecte des données)* décrit les activités d'assurance qualité à réaliser une fois que les activités de terrain sont terminées et que le fichier de données brutes est prêt à être nettoyé. La première partie présente les procédures de nettoyage et de préparation des données à réaliser en vue des calculs de pondération de l'échantillon. La deuxième partie concerne l'évaluation de la qualité de l'échantillonnage, de l'erreur d'échantillonnage et de la pondération de l'échantillon, tandis que la troisième partie se concentre sur l'évaluation de la qualité de la couverture, des taux de non-réponse et d'autres erreurs non liées à l'échantillonnage. Enfin, la quatrième partie décrit la création du fichier de données analytiques final.

Ces chapitres visent à fournir une série de directives et de recommandations fondées sur les meilleures pratiques d'assurance qualité. Les méthodes et documents détaillés relatifs à chaque partie des chapitres sont fournis en annexe, le cas échéant.

1.3 Assurance qualité de l'enquête GATS

L'assurance qualité est un processus qui consiste en des activités systématiques destinées à garantir, évaluer et confirmer la qualité des données recueillies au cours d'une enquête (Biemer et Lyberg, 2003). Si la qualité des données était autrefois synonyme d'exactitude des données, ce terme revêt désormais une acceptation plus large. Ainsi, des données de haute qualité sont aujourd'hui des données qui sont adaptées à leur usage prédestiné. Cette définition implique non seulement l'exactitude, mais aussi des attributs tels que l'actualité, l'accessibilité et la comparabilité des données, conférant ainsi de multiples dimensions au concept de qualité. Un ensemble de données est jugé exact s'il est exempt d'erreurs. Il est jugé opportun s'il est disponible au moment où l'on en a besoin. L'accessibilité est déterminée selon le degré relatif de facilité ou de difficulté d'utilisation. Des données de qualité sont comparables si elles sont identiques d'une unité à l'autre, que la comparaison unitaire s'applique entre individus, enquêteurs et unités primaires d'échantillonnage (UPE) ou à l'échelle internationale. La normalisation augmente la comparabilité des données d'enquête. Par exemple, l'utilisation de questions normalisées garantit que différents enquêteurs poseront tous les mêmes questions, de la même manière. De même, la

normalisation des processus d'assurance qualité permet de garantir que les erreurs de données sont résolues et corrigées de la même manière.

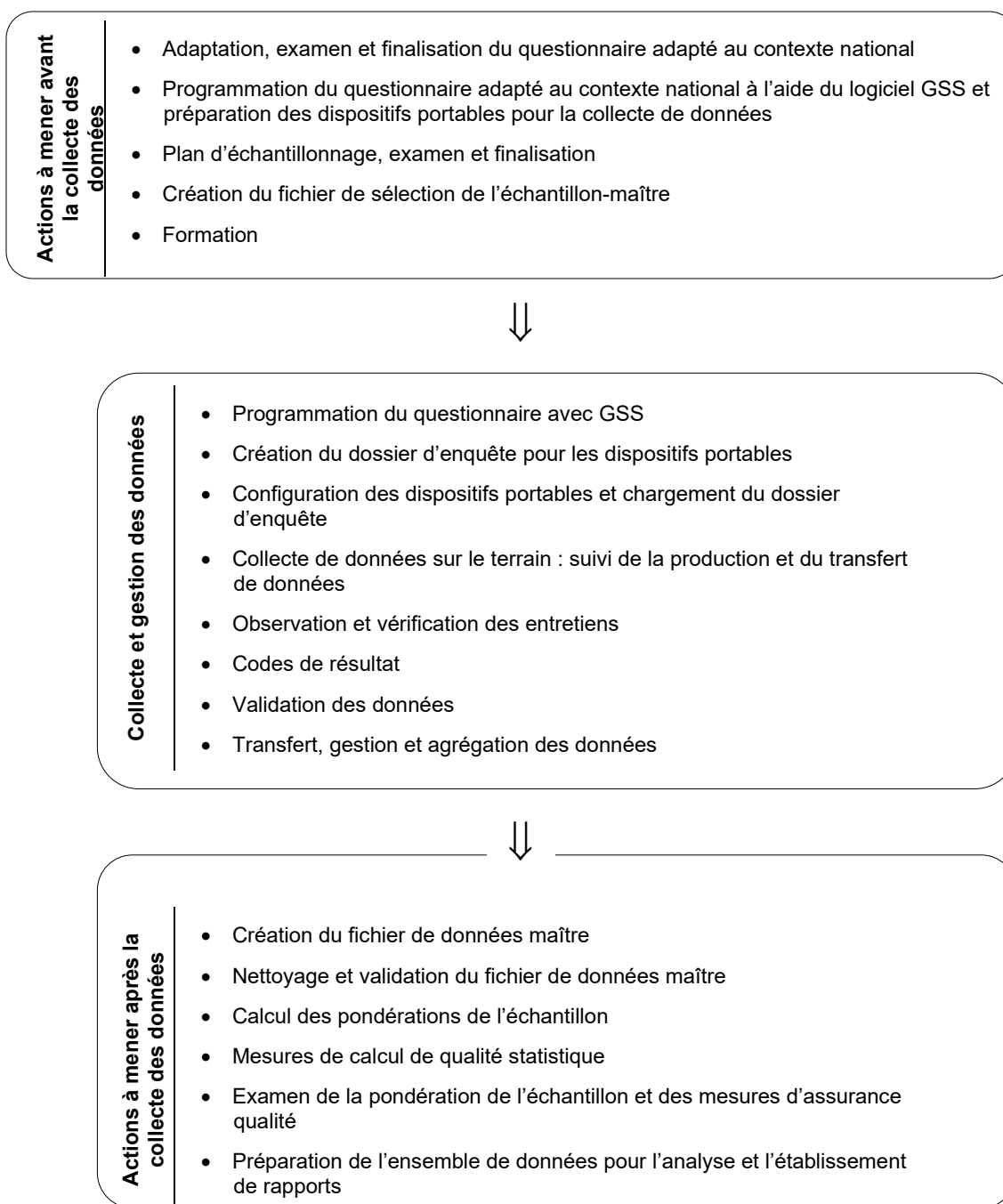
Les multiples dimensions que recouvre le concept de qualité se restreignent souvent mutuellement, de sorte que l'optimisation de la qualité est une question de compromis et d'équilibre entre toutes ces dimensions. Par exemple, des données de qualité doivent non seulement être exactes, mais aussi disponibles en temps voulu. S'il faut beaucoup de temps pour garantir un haut degré d'exactitude, alors les données peuvent ne plus être pertinentes.

Les directives présentées dans ce document constituent des procédures normalisées aux fins de l'assurance qualité des données collectées via la mise en œuvre de l'enquête GATS. Les pays sont encouragés à intégrer des activités d'assurance qualité supplémentaires, en complément de celles décrites dans le présent document.

2. Plan du processus d'assurance qualité

Le document suivant fournit une représentation graphique des activités d'assurance qualité sous forme de calendrier global et de séries d'actions non concomitantes. Ces actions se répartissent en trois phases : la phase préalable à la collecte de données, la phase de collecte de données proprement dite, et la phase postérieure à la collecte de données.

Document 2-1. Diagramme des flux relatifs à l'assurance qualité



3. Assurance qualité : Actions à mener avant la collecte des données

La phase préalable à la collecte des données comprend une série de tâches qui doivent être réalisées afin de préparer la collecte complète des données de l'enquête. Parmi ces tâches figurent la conception et la programmation du questionnaire, la réalisation du test préliminaire, le plan d'échantillonnage, la préparation du fichier de sélection de l'échantillon-maître et la formation. Ce chapitre décrit le mécanisme d'assurance qualité ainsi que les directives normalisées qui doivent être suivies dans ces différents domaines. Ce chapitre traite spécifiquement des points suivants :

- 3.1 Conception du questionnaire et processus d'examen
 - 3.1.1 Orientations relatives à l'adaptation du questionnaire
- 3.2 Programmation et contrôle de versions
 - 3.2.1 Procédures de programmation
 - 3.2.2 Vérification du questionnaire
 - 3.2.3 Dispositifs portables
- 3.3 Test préliminaire
- 3.4 Finalisation du questionnaire
- 3.5 Transfert et gestion des données
- 3.6 Plan d'échantillonnage et processus d'examen
- 3.7 Préparation du fichier de sélection de l'échantillon-maître
- 3.8 Ateliers de formation

3.1 Conception du questionnaire et processus d'examen

L'enquête GATS applique un processus normalisé qui permet aux pays participants de concevoir des questionnaires adaptés à leur contexte national. Le comité d'examen du questionnaire GATS, qui réunit des experts de la lutte antitabac et de la conception de questionnaires issus de divers pays développés et en développement, vérifie et approuve tous les questionnaires GATS. En collaboration étroite avec chaque pays, le comité d'examen veille à adapter le questionnaire GATS au contexte national de chacun, tout en conservant les questions principales normalisées de l'enquête GATS (voir le *Questionnaire principal de l'enquête GATS avec questions facultatives* pour plus de détails) afin de garantir la comparabilité des données entre pays. Une description détaillée des modalités d'adaptation du questionnaire et du processus d'examen est fournie en *annexe B.1*.

Le comité d'examen du questionnaire GATS vérifie et approuve le questionnaire à des fins d'assurance qualité et d'harmonisation entre pays.

3.1.1 Orientations relatives à l'adaptation du questionnaire

Dans le cadre de l'adaptation du *Questionnaire principal de l'enquête GATS*, le comité d'examen du questionnaire recommande aux pays d'appliquer certaines directives spécifiques afin de faciliter le processus d'examen et de veiller au respect des normes de qualité. Ces directives sont les suivantes :

- Surligner toutes les adaptations du *Questionnaire principal de l'enquête GATS* au contexte national (pour en faciliter la consultation).
Cela inclut les listes de questions, les catégories de réponses, les questions facultatives et les questions ajoutées pour le pays concerné.
- Barrer le texte des questions principales que le pays souhaite supprimer (pour en faciliter la consultation).
- Conserver la numérotation et l'ordre des questions principales et facultatives, afin d'assurer une cohérence permettant les comparaisons entre pays. (Même si cela n'est pas toujours possible ni optimal, il convient d'appliquer cette directive autant que possible.)
- Pour les questions nouvellement ajoutées adaptées au contexte national, doubler la lettre correspondant à la section (p. ex., AA10, BB17, EE4). Ainsi, l'ajout de nouvelles questions spécifiques ne perturbera pas la numérotation d'origine.
- Ne modifiez les instructions « passer à » que pour tenir compte des ajouts de questions facultatives ou de questions propres au pays.
- Ne modifiez pas les questions principales normalisées (sauf pour les catégories spécifiques au pays), afin de maintenir la cohérence nécessaire à la comparaison entre pays. (En cas de dérogation à la règle, une justification doit être fournie au comité d'examen du questionnaire.)
- Limiter le nombre de questions supplémentaires afin que le questionnaire conserve une longueur raisonnable.
- Limiter la complexité des questions supplémentaires pour faciliter la programmation du questionnaire dans chaque pays.
- Fournir une justification en cas d'adaptation. Ces explications permettent d'éclairer le comité d'examen du questionnaire lors de l'examen et de l'approbation des questionnaires nationaux.

3.2 Programmation et contrôle de versions

L'enquête GATS s'appuie sur une méthode de collecte de données fondée sur des entretiens individuels assistés par ordinateur, à l'aide de dispositifs portables utilisés comme équipements électroniques. Les dispositifs portables sont équipés du logiciel General Survey System (GSS)¹. Chaque pays présente des caractéristiques qui lui sont propres, s'agissant de l'adaptation de son questionnaire ou de ses langues et systèmes d'écriture, sans oublier les équipements utilisés (par exemple, les modèles d'appareils portables utilisés peuvent différer). Ces facteurs de complexité justifient l'application d'un certain nombre de mesures et de meilleures pratiques.

Des processus et des vérifications standard doivent être appliqués afin de s'assurer de la qualité et de l'efficacité de la programmation des

Il existe une série de mesures (précisées à l'*annexe B.2*) qui requièrent une coopération et une contribution actives du pays hôte longtemps à l'avance (idéalement 6 à 8 semaines avant le début de la formation au test préliminaire). Ces mesures supposent que le pays hôte identifie différents profils

¹ Le General Survey System (GSS) est un logiciel destiné aux appareils Android, qui a été testé et utilisé sur des tablettes Android.

(informaticiens/programmeurs, personnel chargé de l'enquête, gestionnaires d'enquête, etc.) qui soient disponibles pour travailler durant l'atelier d'initiation ou peu après.

Afin d'amorcer le travail préparatoire, ces personnes seront formées en face à face ou au moyen de webinaires. Il est attendu d'elles qu'elles consacrent un nombre d'heures significatif à la préparation au test préliminaire de l'enquête GATS et à la mise en œuvre complète du questionnaire.

Les sous-sections suivantes reviennent sur différents aspects du questionnaire et du processus de préparation de l'enquête auxquels il convient de prêter attention afin d'améliorer la qualité.

3.2.1 Procédures de programmation

Les procédures suivantes doivent être appliquées afin de garantir la qualité et l'efficacité des préparatifs de l'enquête GATS :

- Il n'appartient pas au personnel informatique de vérifier la qualité de la traduction du questionnaire national dans la langue du pays : cette responsabilité incombe au personnel chargé de l'enquête, au personnel de terrain et aux personnes référentes chargées de la mise en œuvre de l'enquête. Le personnel informatique aidera le personnel chargé de l'enquête, le personnel de terrain et les personnes référentes chargées de la mise en œuvre de l'enquête à intégrer le texte du questionnaire dans le logiciel GSS utilisé pour la conception de l'enquête.
- La personne référente de l'organisme de mise en œuvre doit approuver le *questionnaire destiné aux ménages* et le *questionnaire individuel* (par écrit) avant la formation du personnel de terrain (en utilisant des numéros de versions).
- Modifications du texte : Toute modification du texte du questionnaire doit être prise en compte dans toutes les langues dans lesquelles le questionnaire sera administré.
- Toute modification de la logique de programmation (fourchettes et pages, vérifications de validité, etc.) d'un questionnaire déjà approuvé doit faire l'objet d'un nouvel examen par le QRC.
- La modification du contenu ou de la formulation d'une question ou des réponses doit faire l'objet d'un nouvel examen par le QRC.
- Contrôle des versions des fichiers DB3 (fichiers de la base de données du questionnaire d'enquête) : Le pays hôte est propriétaire des fichiers DB3, et le personnel informatique/chargé de l'enquête au niveau national doit conserver une trace des modifications apportées aux questionnaires ainsi que des spécifications de programmation dans GSS.
- Après le bilan du test préliminaire et le démarrage de la mise en œuvre complète de l'enquête, le pays doit envoyer les fichiers DB3 finaux CMSDB, Survey0 et Survey1.

3.2.2 Vérification du programme relatif au questionnaire

Il existe plusieurs mesures d'assurance qualité à appliquer et à documenter au titre de la vérification du programme relatif au questionnaire :

- Les responsables de la programmation du questionnaire doivent mettre à jour la dernière version à chaque fois qu'une modification y est apportée. Le numéro de version doit par ailleurs être confirmé avant l'examen du programme relatif au questionnaire. Une archive des fichiers DB3 plus anciens est conservée dans le logiciel GSS à titre d'historique et de sauvegarde.
- Pour chaque langue, les spécifications de programmation complètes sur GSS sont examinées en visualisant chaque question sur le dispositif portable et en vérifiant que le texte de la question, les réponses et les caractères spéciaux correspondent exactement à ceux du document *Spécifications pour la programmation du questionnaire GATS*. Si des problèmes sont constatés, ils sont clairement documentés sur les spécifications ; après une itération, les problèmes sont corrigés, puis l'itération de vérification complète est effectuée. Ce processus itératif se poursuit jusqu'à ce que toutes les questions aient été approuvées sur le dispositif portable. Cette procédure doit absolument être respectée à la lettre, et réalisée avec l'aide d'un expert en enquêtes parlant la langue locale. Elle ne doit pas être confiée aux programmeurs, ni à du

personnel ne possédant pas l'expertise linguistique requise. La personne référente ou représentante au niveau du pays doit valider la version finale après examen des modifications/mises à jour réalisées. Les spécifications doivent être passées en revue avant validation afin de s'assurer que chaque question a été vérifiée en se référant au questionnaire (pour plus de détails, voir le *module 2.1* du *Guide de formation à la gestion des données de l'enquête GATS*).

- Les instructions « passer à » et le texte en anglais sont toujours testés avant la visite sur site. Le logiciel d'enquête ne doit pas être modifié durant la collecte de données sur le terrain, à moins que cela ne soit absolument nécessaire.
- Le nombre de balises HTML ouvrantes et fermantes et de symboles (,), [,], { , et } doit être identique dans chaque langue. Dans le cas contraire, le document doit être revérifié et le problème résolu.

3.2.3 Dispositifs portables

Il existe des procédures formelles d'assurance qualité à suivre concernant le chargement du programme GSS sur les dispositifs portables :

- Avant le chargement, tous les dispositifs portables doivent avoir terminé la configuration initiale et, le cas échéant, les versions précédentes du logiciel GATS doivent avoir été désinstallées et les dossiers afférents supprimés.
- Des étiquettes indiquant la liste des actions de contrôle qualité à remplir (pour l'installation de l'enquête GATS sur les dispositifs portables) sont imprimées et apposées sur les dispositifs portables. La personne qui effectue le contrôle qualité sur le dispositif portable doit s'assurer que chaque étape a été accomplie.

L'*annexe B.3* décrit plus en détail les différentes étapes à suivre pour le chargement des dispositifs portables (pour plus de détails, se référer au *module 2.2.* du *Guide de formation à la gestion des données de l'enquête GATS*).

3.3 Test préliminaire

La réalisation d'un test préliminaire à l'échelle nationale en amont du déploiement complet de l'enquête répond à six grands objectifs. L'atteinte des objectifs suivants contribuera à garantir la qualité de la collecte de données de l'enquête GATS complète :

1. Fournir au personnel du pays une assistance en matière de programmation, d'essai et de poursuite d'une enquête au moyen de dispositifs portables, selon les besoins.
2. Examiner le questionnaire du test préliminaire de l'enquête GATS, y compris la formulation des questions et des réponses, les demandes d'approfondissement, la configuration du branchement conditionnel, ainsi que la durée et le moment de l'entretien.

Le test préliminaire de l'enquête GATS permet de mettre à l'essai le questionnaire et les procédures de collecte et de gestion des données afin de s'assurer que l'enquête complète sera menée suivant

3. Fournir au personnel du pays une assistance en matière d'utilisation des dispositifs portables et des logiciels embarqués, selon les besoins.
4. Observer la mise en œuvre du test préliminaire en milieu rural et en milieu urbain, si cela s'avère possible et pertinent.
5. Tester la collecte, le transfert et le plan de gestion des données pendant le test préliminaire, afin d'identifier les adaptations à prévoir en vue de la mise en œuvre complète de l'enquête.
6. Présenter aux partenaires de l'enquête GATS un bilan des forces, des faiblesses et des recommandations formulées concernant la préparation et la mise en œuvre du test préliminaire.

3.4 Finalisation du questionnaire

Afin de finaliser le questionnaire GATS en vue de la mise en œuvre complète de l'enquête, les étapes à remplir sont les suivantes :

- **Bilan du test préliminaire.** Tous les membres du personnel (enquêteurs de terrain, superviseurs de terrain, équipe de projet) doivent donner leur avis sur le déroulement du questionnaire dans son ensemble, ainsi que sur les questions d'enquête spécifiques à cibler pour les révisions.
- **Examen des données du test préliminaire.** L'analyse des données du test préliminaire permet de mettre en lumière les ajustements à réaliser à l'issue du test préliminaire. Voici quelques problèmes spécifiques à examiner :
 - Non-réponse à une question : Identifiez les questions présentant les plus hauts niveaux de réponses « Ne sait pas » ou « N'a pas souhaité répondre ».
 - Fréquences : Examinez les répartitions de fréquence des questions principales afin de déterminer si les données semblent appropriées. Toute anomalie peut être le signe d'un problème dans la conception du questionnaire ou autre (problème de formation, par exemple).
 - Plages admises : Vérifier la pertinence des plages admises programmées sur les dispositifs portables et les adapter à la mise en œuvre complète de l'enquête, le cas échéant.
 - Catégories de réponses : Déterminer s'il était opportun d'adapter le questionnaire au contexte national (par exemple, en modifiant les listes de questions et les catégories de réponses). Par exemple, mesurer la fréquence des marques de cigarettes mentionnées afin de modifier la liste de marques dans la version finale du questionnaire.
- **Révision de la mise en œuvre complète de l'enquête.** La version révisée à la suite du test préliminaire doit être soumise à l'examen du Comité d'examen du questionnaire, avec une justification des révisions (en fonction des résultats du test préliminaire). Les mêmes directives d'assurance qualité que celles décrites dans les *sections 4.1 et 4.2* (selon le cas) doivent être suivies dans le cadre de la préparation du questionnaire pour la mise en œuvre de la version finale de l'enquête (y compris la programmation révisée, etc.).

3.5 Transfert et gestion des données

Le test préliminaire consiste en une mise à l'essai complète du déploiement de la version finale de l'enquête pour le processus de gestion des données. À la lumière de l'expérience et des recommandations du test préliminaire, le processus de transfert et de gestion des données sera finalisé en vue du déploiement de la version finale de l'enquête. Par conséquent, chaque étape du processus de gestion des données doit être menée conformément au plan de gestion des données, en tenant compte de l'enquête complète. Par exemple :

- Le mécanisme de retour d'informations concernant le partage de rapports doit être testé lors du test préliminaire.
- La méthode de transmission des données (par WiFi, SIM, cloud ou site FTP) doit être mise à l'essai pendant le test préliminaire et les ajustements nécessaires, le cas échéant, doivent être effectués pour la version finale en fonction des enseignements tirés du test préliminaire.

3.6 Plan d'échantillonnage et processus d'examen

L'enquête GATS applique un processus normalisé qui permet aux pays participants d'élaborer leur propre plan d'échantillonnage. Le Comité d'examen des échantillons de l'enquête GATS, un groupe composé d'experts en méthodes d'échantillonnage issus de pays développés et en développement, examine et finalise les plans d'échantillonnage des pays participants. Les personnes référentes du Comité d'examen des échantillons et des Centers for Disease Control and Prevention (CDC) travaillent en étroite collaboration avec chaque pays pour adapter le plan d'échantillonnage de l'enquête GATS tout en respectant les normes GATS afin de garantir la comparabilité des données entre pays. Une description détaillée du plan d'échantillonnage de l'enquête GATS est fournie dans le *Manuel d'échantillonnage de l'enquête GATS*.

Dans le cadre du processus d'échantillonnage de l'enquête GATS, il peut être nécessaire de réaliser une cartographie et de constituer des listes afin de créer une base d'échantillonnage aléatoire de surface appropriée. Pour consulter les procédures et les recommandations complètes d'assurance qualité en matière de cartographie et de listes, se référer au *Manuel de cartographie et de constitution des listes GATS*.

3.7 Préparation du fichier de sélection de l'échantillon-maître

Une fois que l'échantillon GATS a été dressé, un fichier de sélection de l'échantillon-maître doit être préparé. Le fichier de sélection de l'échantillon-maître est un ensemble de données qui contient un numéro d'enquête pour chaque ménage de la liste d'échantillonnage, ainsi que les informations nécessaires pour calculer les pondérations d'échantillon et analyser les données complexes de l'enquête.

Le fichier de sélection de l'échantillon-maître inclut des identifiants d'échantillon et des identifiants géographiques. Les identifiants de l'échantillon tels que la strate, l'unité primaire d'échantillonnage (UPE), l'unité secondaire d'échantillonnage (USE/segment), le type de lieu de résidence (urbain/rural) et la pré-attribution du genre du ménage, si une randomisation selon le genre est utilisée. Informations de géolocalisation, telles que les

Le fichier de sélection de l'échantillon-maître comporte un numéro d'enquête (CaseID) unique, des identifiants d'échantillon, ainsi que des

coordonnées avec les codes du pays, de la région, de la province/l'État, du comté/district, ainsi que de la ville ou du village.

Une fois le fichier de sélection de l'échantillon-maître finalisé, celui-ci doit être organisé et converti en fichier texte en utilisant Excel et/ou Access. Ce fichier texte sera utilisé pour charger les enquêtes sur les dispositifs portables destinés à la collecte de données, le cas échéant. De plus, en utilisant le logiciel GSS, les écrans Case Grid et Case Info doivent être finalisés pour refléter le dossier d'enquête révisé. Le chapitre suivant fournit de plus amples détails au sujet de ce processus.

3.8 Ateliers de formation

La formation constitue un important mécanisme d'assurance qualité dans le cadre de la mise en œuvre de l'enquête GATS. Les ateliers de formation dispensés à l'intention des pays couvrent l'introduction à l'enquête GATS, le test préliminaire (abordé dans les précédentes sections) ainsi que l'enquête complète. Ces formations font partie du protocole et des normes de l'enquête GATS permettant d'assurer la collecte de données de qualité pour chaque pays.

Les formations à l'enquête GATS sont dispensées à tous les niveaux de l'équipe nationale d'enquête impliquée dans la mise en œuvre de l'enquête, des enquêteurs aux superviseurs, en passant par le personnel informatique et l'équipe centrale qui supervise l'ensemble du processus de bout en bout. Des partenaires nationaux et internationaux assurent la formation à l'enquête GATS à tous les niveaux. La formation de l'ensemble du personnel permet de s'assurer que toutes les personnes impliquées dans le processus d'enquête connaissent parfaitement leur rôle pour garantir la qualité des données. La formation relève d'un processus continu, mené à la fois avant et pendant la mise en œuvre de l'enquête. Chaque session de formation se termine par la réalisation d'un bilan complet.

Les objectifs de formation relatifs à l'enquête GATS sont les suivants :

- Veiller à l'application normalisée du protocole, des normes et des supports d'enquête ;
- S'assurer que l'objectif de l'étude et le protocole d'enquête sont bien compris.
- S'assurer de l'application standard du système de collecte, de gestion, de transfert et d'agrégation des données électroniques de l'enquête GATS.
- Impliquer et motiver les enquêteurs ;
- Dispenser des conseils pratiques ;
- Rehausser la qualité globale des données produites.

3.8.1 Atelier d'introduction à l'enquête GATS

L'atelier d'introduction constitue la première étape d'assurance qualité du processus GATS. Cela implique un examen approfondi du protocole normalisé de l'enquête GATS, y compris le questionnaire, le plan d'échantillonnage, la collecte et la gestion des données électroniques, ainsi que la diffusion. Les pays partagent leurs expériences en matière de réalisation d'enquêtes, afin de démontrer leur capacité technique à mettre en œuvre l'enquête GATS dans le pays, mais aussi à déterminer l'assistance technique requise de la part de partenaires de l'organisme afin d'atteindre les résultats souhaités. Les ateliers d'introduction passent en revue les directives relatives à l'élaboration de propositions et permettent aux pays d'adapter le protocole à leur propre situation. Enfin, ils mettent à disposition un

guide étape par étape pour l'élaboration, la soumission, l'examen et le financement de la proposition, outil indispensable au lancement et à la mise en œuvre du projet d'enquête.

3.8.2 Test préliminaire

Les objectifs de la réalisation d'un test préliminaire ont été exposés dans la *section 3.3*. La formation au test préliminaire comporte de multiples composantes, dont la formation du personnel informatique, la formation des enquêteurs/superviseurs sur le terrain, la pratique sur le terrain et les modalités de bilan. Le test préliminaire permet à l'équipe de projet d'accumuler de l'expérience en matière de formation du personnel de terrain. Au cours du test préliminaire, le personnel informatique et de terrain est formé aux logiciels et à l'équipement nécessaires à la mise en œuvre de l'enquête GATS. La formation permet de s'assurer que le personnel est correctement préparé et comprend le fonctionnement des logiciels et du matériel nécessaires à la mise en œuvre de l'enquête GATS.

3.8.3 Formation complète sur l'enquête

La formation des enquêteurs et des superviseurs à la mise en œuvre de l'enquête complète constitue l'un des principaux mécanismes d'assurance qualité de l'enquête GATS. L'objectif est de former les enquêteurs et les superviseurs à la production de données utiles et fiables à partir de l'enquête de terrain. La formation couvre les dimensions suivantes : raison d'être de l'enquête, contexte de la lutte antitabac et système mondial de surveillance du tabac, techniques d'entretien, identification des ménages échantillonnés, consentement, contenu du questionnaire et éléments pratiques, notamment jeux de rôle et observation sur le terrain. Les formateurs passent en revue le questionnaire papier ou le questionnaire sur tablette électronique avec les participants, afin de les aider à mieux comprendre l'enquête. En outre, par le biais de jeux de rôle et d'entraînement pratique sur le terrain, les stagiaires se familiarisent avec le questionnaire et avec les problèmes pouvant survenir lors d'un entretien avec des répondants difficiles, ainsi que les approches à adopter pour gérer ces situations. Des supports de formation sont également fournis aux personnes formées à titre de matériel de référence, durant la formation et le travail de terrain. Parmi ces supports figurent le questionnaire, les spécifications question par question, ainsi que le manuel d'enquête.

4. Assurance qualité : Collecte et gestion des données

Ce chapitre décrit les processus et procédures d'assurance qualité des données appliqués dans le cadre de la collecte et de la gestion des données de l'enquête GATS. La collecte et la gestion des données englobent la configuration du logiciel GSS de l'enquête GATS, l'utilisation du logiciel pendant la collecte de données sur le terrain, le transfert des données collectées, ainsi que l'agrégation et la création du fichier maître de la base de données agrégée. Elles concernent également le suivi, l'établissement de rapports et l'appui technique durant la collecte de données sur le terrain.

Pour tout pays donné, la collecte, l'agrégation et l'assurance qualité des données doivent être effectuées à deux niveaux : sur le terrain, d'une part, et au niveau du Centre national des données du pays, de l'autre. Ce chapitre décrit les processus et les procédures d'assurance qualité des données applicables à chacun de ces deux niveaux, en abordant spécifiquement les points suivants :

4.1 Collecte de données sur le terrain : Configuration et maintenance

4.1.1 Dossier d'enquête et gestion des enquêtes

4.1.2 Dispositifs portables de collecte de données électroniques et logiciel d'enquête

4.1.3 Procédures d'assurance qualité de la collecte de données sur le terrain

4.2 Gestion des données du Centre national des données

4.2.1 Agrégation

4.2.2 Contrôle qualité des données collectées

4.2.3 Communication et appui technique

4.1 Collecte de données sur le terrain : Configuration et maintenance

Le protocole standard pour l'enquête GATS recommande que certaines procédures systématiques soient établies et suivies par les employés de terrain et le personnel informatique, afin de garantir l'exactitude et l'exhaustivité des données tout au long de la période de collecte de données. Ces procédures sont décrites ci-dessous.

4.1.1 Dossier d'enquête et gestion des enquêtes

1. Il est recommandé que le dossier d'enquête soit préparé et tenu à jour au niveau du Centre national des données en utilisant le logiciel GSS, le cas échéant, selon le plan de gestion des données du pays. (Notez que le dossier d'enquête constitue un sous-ensemble du fichier de sélection de l'échantillon-maître).
2. Durant la création du dossier d'enquête et avant la configuration des dispositifs portables :
 - Tous les enquêteurs de terrain doivent être identifiés et se voir attribuer un dispositif portable. Les numéros d'enquête attribués doivent être uniques et spécifiques à chaque enquêteur de terrain, en utilisant la variable correspondante dans le dossier d'enquête. Au moment de la configuration, un même numéro d'enquête ne doit pas être assigné à plus d'un enquêteur de terrain. Cette précaution permet d'éviter les doublons, tout en s'assurant que le nombre d'enquêtes assignées à chaque enquêteur de terrain est raisonnable et que les enquêtes sont facilement identifiables.

- Tous les enquêteurs de terrain doivent posséder un identifiant numérique à 6 chiffres, conformément à la convention GATS.
- Tous les numéros d'enquête doivent être un identifiant numérique à 6 chiffres, conformément à la convention GATS.

3. Le dossier d'enquête doit comporter un identifiant pour chaque dispositif portable et enquêteur de terrain qui collectera des données.
 - Assurez-vous que tous les enquêteurs de terrain inscrits sur le tableur d'identification des enquêteurs de terrain sont bien répertoriés dans le dossier d'enquête. Le fichier doit être examiné avant la configuration des dispositifs portables. Une fois les enquêtes chargées sur les dispositifs portables, le nombre escompté d'enquêtes doit apparaître dans le système de gestion des enquêtes. Si ce n'est pas le cas, vérifiez l'identifiant d'enquêteur de terrain attribué dans le logiciel du dispositif portable et l'identifiant figurant dans le dossier d'enquête. Notez que chaque numéro d'enquête donnera lieu à la création de deux registres qui seront chargés sur le système de gestion des enquêtes (HQ=00 et IQ=01).
4. La grille d'enquêtes sur chaque dispositif portable contiendra uniquement les enquêtes attribuées et confiées à l'enquêteur de terrain concerné. En général, 300 numéros d'enquête au maximum doivent être attribués à chaque enquêteur de terrain et chargés dans le dispositif portable. Les performances, telles que le temps de chargement de l'application et la possibilité d'identifier facilement une enquête pendant le travail sur le terrain, peuvent se trouver affaiblies si le nombre d'enquêtes chargées dépasse le seuil recommandé.
5. Avant de procéder à l'attribution des enquêtes sur chaque dispositif portable, effectuez un examen des fréquences des champs du dossier d'enquête sur l'ordinateur du Centre national des données. Pour connaître les exigences relatives aux champs, référez-vous au document de dossier d'enquête sur le système de gestion des enquêtes (*Guide de programmation sur General Survey System*) :
 - Vérifiez que le champ CASEID ne comporte pas de doublons et qu'il se compose bien d'un identifiant numérique à 6 chiffres, conformément à la norme GATS.
 - Vérifiez que le champ PROJECTNAME indique le nom du pays, afin de s'assurer qu'il existe un unique champ d'identification du pays dans le dossier d'enquête (la requête de mise à jour du champ « projectname » peut être utilisée dans le modèle de dossier d'enquête sur Access).
 - Vérifiez que la valeur de la variable CreateDate correspond bien à la date de création du fichier.
 - Comparez le fichier de sélection de l'échantillon-maître au dossier d'enquête sur les dispositifs portables, afin d'en vérifier l'exhaustivité et l'exactitude. (Notez que le dossier d'enquête sur dispositif portable constitue un sous-ensemble du fichier de sélection de l'échantillon-maître).

Le *Guide de programmation sur General Survey System* contient des informations détaillées au sujet de la création du dossier d'enquête, des descriptions des champs et du tableur d'inventaire des enquêteurs de terrain sur les dispositifs portables.

4.1.2 Dispositifs portables de collecte de données électroniques et logiciel d'enquête

1. **Paramétrage des informations relatives aux enquêteurs de terrain** : Chaque dispositif portable doit être étiqueté avec un identifiant d'enquêteur de terrain unique suivant la convention GATS, avec le même identifiant défini et affiché dans le menu Paramètres du système de gestion des enquêtes. Ces informations doivent être enregistrées dans un tableur conservé au Centre

national des données, avec le nom de l'enquêteur de terrain et le numéro de série du dispositif portable.

- Recoupement des paramètres d'identification des enquêteurs de terrain : avant la collecte de données, vérifiez que le tableur attribué, le champ d'identification de l'enquêteur de terrain dans le menu Paramètres et l'étiquette du dispositif portable sont tous identiques.
2. **Versions des logiciels : système de gestion des enquêtes, questionnaire destiné aux ménages et questionnaire individuel** : Les dispositifs portables doivent contenir la version finale approuvée du questionnaire.
- Pour chaque dispositif portable, vérifiez que la version de chaque questionnaire est bien chargée et fonctionnelle en cochant l'option « Information » en haut de l'écran et en vérifiant que les versions dans la section « Informations d'enquête » sont correctes.
 - Dans la mesure du possible, tous les dispositifs portables doivent être configurés en même temps afin de garantir la cohérence des versions, de la configuration et du processus de contrôle qualité.
3. **Chargement des enquêtes** : Au début de la collecte de données/du travail sur le terrain, seules les enquêtes répertoriées dans le dossier d'enquête avec un identifiant d'enquêteur de terrain correspondant doivent apparaître dans le système de gestion des enquêtes de chaque dispositif portable. Les enquêtes utilisées pour la formation doivent être effacées avant le démarrage du travail de terrain, afin d'éviter tout risque de travailler par inadvertance sur une enquête utilisée lors de la formation et de l'exporter. Dans les cas où des enquêteurs de terrain ne respecteraient pas cette règle, les analystes doivent se préparer à filtrer les enquêtes utilisées lors de la formation afin de ne pas les intégrer aux données réelles.

Le *Guide de programmation sur General Survey System* contient des instructions détaillées concernant le paramétrage des dispositifs portables ainsi que la création et l'installation des fichiers relatifs aux dispositifs portables.

4.1.3 Procédures d'assurance qualité de la collecte de données sur le terrain

Cette section offre un aperçu des procédures d'assurance qualité recommandées qui peuvent être mises en œuvre pendant la collecte de données sur le terrain. Des informations plus détaillées sont disponibles dans le *Manuel d'enquête sur le terrain* et dans le *Manuel de supervision sur le terrain* de l'enquête GATS.

Rôle du superviseur de terrain

Les superviseurs de terrain assurent l'interface entre les enquêteurs de terrain et le personnel de gestion de l'enquête GATS. Il leur appartient d'assurer le suivi de la production et des résultats, ainsi que de communiquer tout problème sur le terrain susceptible de retarder l'achèvement de l'enquête GATS par rapport aux délais prévus. Le superviseur est chargé de veiller à ce que les données de leur équipe soient collectées conformément au protocole prévu à cet effet et à ce qu'elles soient de qualité optimale. Pour que les enquêteurs de terrain saisissent l'importance que l'équipe GATS accorde à la qualité des données, le superviseur de terrain doit inscrire le contrôle de la qualité dans le cadre de ses activités hebdomadaires.

La première action à entreprendre par le superviseur de terrain pour garantir la qualité des données consiste à informer ses enquêteurs de terrain des niveaux de qualité escomptés. Les superviseurs sont censés fournir aux enquêteurs de terrain un retour d'informations continu sur les questions de qualité (aussi bien positif que négatif) et s'attacheront à souligner l'importance que l'équipe GATS accorde à la *qualité*.

Suivi de la production

La production inclut toutes les activités nécessaires à l'atteinte des objectifs de taux de réponse de l'enquête GATS. Ces objectifs incluent le lancement des activités pour chaque ménage attribué, la prise de contact, la localisation et la conversion des refus, ainsi que la réussite de ces activités, en accord avec toutes les spécifications de l'enquête GATS. Le superviseur de terrain et le personnel chargé de l'enquête doivent travailler de concert afin de chronométrer le temps nécessaire à la réalisation d'un entretien sur une base régulière. En outre, le suivi de la production pendant la collecte des données est essentiel pour s'assurer que l'enquête GATS est conduite à son terme dans les délais et dans les limites du budget alloué.

Le logiciel GSS permet de générer les types de rapports suivants, qui doivent être utilisés par le personnel chargé de l'enquête ou le personnel informatique du pays pour suivre les différents aspects de la production. Idéalement, ces rapports doivent également être communiqués aux superviseurs de terrain sur une base régulière. Le moment et l'endroit où ces rapports seront disponibles dépendront du modèle de gestion des données en place dans chaque pays.

- *Statut des enquêtes du questionnaire destiné aux ménages et du questionnaire individuel* Ces rapports fournissent la liste des codes de résultat les plus récents pour les enquêtes auprès des ménages et les enquêtes individuelles, classés par enquêteur. Ils sont consultés au niveau national et peuvent être communiqués aux superviseurs de terrain et/ou aux enquêteurs de terrain.
- *Rapport d'exceptions* : Ce rapport peut aider le personnel de l'enquête GATS à suivre le nombre de secondes visites pour toutes les enquêtes qui seraient incomplètes et à comprendre les raisons motivant ces nouvelles visites.
- *Rapport de répartition des fréquences pour le questionnaire individuel* : Ce rapport peut aider le personnel chargé de l'enquête GATS et le personnel informatique à suivre les réponses au questionnaire par enquêteur ou par UPE, afin d'évaluer la qualité des entretiens de l'enquête en analysant divers indicateurs de qualité :
 - Problèmes potentiels liés à l'enchaînement des entretiens ou aux données « manquantes ».
 - Questions présentant des niveaux de réponses « Ne sait pas » ou « N'a pas souhaité répondre » plus élevés que prévu.
 - Taux de réponses « Autre » plus élevés que prévu (par rapport aux options de réponse prédéfinies).
 - Réponses « Autre, préciser » manquant de clarté, difficiles à coder ou en doublon avec une option prédéfinie dans la liste.

Suivi du transfert de données

La gestion de l'enquête GATS suppose une transmission régulière des données des enquêteurs de terrain au Centre national des données, afin que tout problème potentiel soit détecté et résolu en amont. Il incombe aux superviseurs de terrain de s'assurer que chaque enquêteur de terrain transmet régulièrement les données collectées (idéalement à un rythme quotidien). Les superviseurs de terrain peuvent travailler avec le personnel informatique pour déterminer s'il existe des problèmes au niveau des transmissions.

Observation

L'observation du travail des enquêteurs de terrain est l'un des principaux volets du contrôle qualité assuré par le superviseur de terrain, afin d'aider lesdits enquêteurs à mieux accomplir leur tâche. L'observation contribuera également à éviter toute tentation de leur part de passer outre certaines étapes dans l'administration du questionnaire ou de ne pas appliquer le protocole relatif aux entretiens de l'enquête GATS. Les superviseurs de terrain sont tenus d'observer chacun de leurs enquêteurs de terrain au moins pendant les premiers jours de l'enquête, puis occasionnellement, suivant les directives de l'organisme de mise en œuvre. Dans le cadre de ces observations, le superviseur de terrain accompagnera l'enquêteur de terrain auprès des ménages de l'échantillon pour s'assurer que celui-ci enregistre correctement le résultat de la visite dans le système de gestion des enquêtes sur GSS et qu'il réalise correctement la sélection du ménage ainsi que l'entretien qui s'ensuit. Ces premières visites contribueront grandement à garantir que les erreurs d'administration du questionnaire ou d'utilisation des codes de résultat sont identifiées rapidement et que l'enquêteur de terrain pourra recevoir une formation ou des conseils supplémentaires si nécessaire.

En outre, les superviseurs de terrain peuvent décider d'accompagner les enquêteurs de terrain lors de visites problématiques (par exemple auprès de ménages ayant précédemment refusé de répondre à l'enquête).

Vérification

L'une des manières de vérifier la qualité des données collectées par le personnel d'enquête consiste à mener de brefs entretiens de vérification auprès de ménages qui ont déjà été sélectionnés et interrogés. Ce court entretien de vérification permettra aux superviseurs de terrain de s'assurer que l'enquêteur de terrain s'est bien acquitté des tâches suivantes :

- Identification et sélection du bon ménage. Il peut arriver qu'un enquêteur de terrain sélectionne un ménage différent de celui prévu dans l'échantillon. Auquel cas, le superviseur de terrain doit demander à l'enquêteur de terrain de se rendre auprès du bon ménage afin de procéder à sa sélection et à l'entretien avec ses membres.
- Consignation correcte de l'âge, du sexe et du statut tabagique des membres du ménage. Si l'âge n'est pas correctement transcrit (par exemple si l'enquêteur de terrain inclut les résidents de moins de 15 ans sur la liste, en plus des résidents de 15 ans ou plus), cette erreur se répercutera sur l'identification des personnes éligibles et l'enquêteur de terrain risquerait de sélectionner un membre du ménage non admissible pour l'entretien.
- Administration du *questionnaire individuel* au membre sélectionné du ménage.

Le nombre exact de ces entretiens de vérification sera déterminé en concertation avec l'équipe de direction de l'organisme de mise en œuvre. Cependant, il est généralement attendu des superviseurs de

terrain qu'ils réalisent des entretiens de vérification aléatoires auprès d'environ 10 % des ménages attribués à chaque enquêteur de terrain. Ces entretiens peuvent impliquer 1) une brève vérification consistant à poser au répondant sélectionné un petit nombre de questions pour confirmer qu'il a récemment répondu à une enquête sur des sujets liés au tabagisme et pour évaluer la performance de l'enquêteur, ou 2) si un pays le souhaite, un nouvel entretien complet basé sur le *questionnaire destiné aux ménages* et, si possible, sur le *questionnaire individuel*. Les répondants étant susceptibles de percevoir le nouvel entretien comme un processus laborieux, il n'est pas nécessaire de leur administrer l'intégralité du *questionnaire individuel*. En outre, les entretiens de vérification peuvent être réalisés sur papier ou via un dispositif portable sur lequel ont été chargés les ménages faisant l'objet d'une vérification. Ne jamais réutiliser le dispositif portable de l'enquêteur de terrain pour conduire les seconds entretiens, mais utiliser un autre dispositif portable.

Le superviseur de terrain comparera ensuite les réponses obtenues lors du second entretien avec celles initialement transcrites par l'enquêteur de terrain. Si vous remarquez des différences et qu'il semble que l'enquêteur de terrain s'est entretenu avec une personne non admissible du ménage, le superviseur de terrain renverra l'enquêteur de terrain sur place pour qu'il interroge la personne pertinente, et les réponses devront ensuite être à nouveau comparées.

Dans tous les cas, qu'une erreur ait été découverte dans le processus de vérification ou non, les informations de l'entretien de vérification doivent être transmises au Centre national des données avec tous les autres supports.

Attribution de codes de résultats définitifs pour les ménages non interrogés

Les superviseurs de terrain doivent approuver le code de résultat définitif pour les ménages non interrogés avant que le questionnaire ne soit finalisé. (Le *Manuel de supervision sur le terrain de l'enquête GATS* fournit des informations supplémentaires pour tenter de terminer les enquêtes en attente auprès de ménages non interrogés). Pour les enquêtes ne nécessitant aucune action supplémentaire, l'enquêteur de terrain doit saisir le code de résultat définitif qui convient pour les ménages non interrogés dans le système de gestion des enquêtes (voir le *Manuel d'enquête sur le terrain* pour la liste et la description des codes de résultat sur le terrain).

Au fur et à mesure de la collecte de données dans une zone donnée (UPE, région, État, etc.), un code de résultat définitif doit être attribué à chacune des enquêtes remplies. Le programme attribuera automatiquement un code de résultat définitif à chaque *questionnaire destiné aux ménages* ou *questionnaire individuel* rempli. Lorsque le *questionnaire destiné aux ménages* ou le *questionnaire individuel* n'ont pas abouti à un entretien complet (par exemple lorsque la personne n'a pas souhaité répondre ou qu'il n'y avait personne au foyer), les enquêteurs de terrain doivent attribuer les codes de résultats définitifs destinés aux ménages non interrogés, avec l'accord de leurs superviseurs.

Au fur et à mesure de la collecte des données, un code de résultat définitif doit être attribué à chacune des enquêtes remplies.

4.2 Gestion des données du Centre national des données

Cette section décrit les procédures de collecte et de gestion des données au niveau du Centre national des données : combinaison/agrégation des fichiers, contrôle qualité (y compris en matière de rapports et de suivi), communication et appui technique, et création de fichiers de données brutes.

4.2.1 Agrégation

Au fur et à mesure de la réception des données au point d'agrégation (serveur en nuage, serveur FTP, etc.), les fichiers doivent être examinés, traités et soumis à un contrôle qualité. Les fichiers de données doivent être documentés et archivés.

Lors du traitement, le logiciel d'agrégation trie et traite les fichiers du dossier sélectionné (par exemple, C:\GATSAndroid2\IncommingDB3s\...) par ordre alphabétique et sans tenir compte de la date. Il est essentiel d'agréger les fichiers dans l'ordre dans lequel ils ont été exportés. Au moment de l'agrégation de tous les fichiers exportés, la convention [nom de l'enquêteur de terrain]#_AAAA_MM_JJ.db3 utilisée pour nommer les fichiers exportés permet un tri automatique par ordre alphabétique, assurant ainsi un traitement adéquat des données. Les fichiers exportés sont cumulatifs, ce qui signifie que le fichier le plus récent contient toutes les données, y compris les mises à jour ou les modifications apportées aux fichiers précédents.

Le traitement de l'agrégation ne permet pas de doublons de numéros d'enquête.

Traitement de l'agrégation

- Au début de la collecte de données, il faut utiliser comme point de départ le fichier d'agrégation maître nommé master.db3zero, disponible dans le dossier C:\GATSAndroid2\DATA. L'agrégation commence ainsi avec un fichier maître ne contenant aucune donnée. Le fichier master.db3zero doit être copié, puis renommé sous l'appellation master.db3 lorsqu'une réinitialisation est nécessaire.
- Un fichier Node.id doit être conservé dans le dossier C:\GATSAndroid2\
 - Avant que l'agrégation ne soit exécutée pour la première fois sur la machine, le fichier de nœuds doit être modifié de façon à avoir un nom unique et pertinent.
 - Une fois l'agrégation réalisée, le fichier Node.ID ne doit plus être modifié.
- Pour les pays réalisant un cryptage des données, les derniers fichiers DB3 collectés doivent être décryptés à l'aide de l'option Agrégation de données > Décrypteur de base de données du logiciel GSS. Si les données sont cryptées, cette action doit être réalisée avant l'agrégation.
- Pour plus d'efficacité, agréger les derniers fichiers DB3 collectés à ceux précédemment agrégés, en s'assurant simplement que les fichiers sont reçus et traités dans l'ordre chronologique (par date d'exportation). S'il est possible d'agréger vers un fichier maître vide à chaque agrégation, il convient de noter que cette méthode prend plus de temps, puisque tous les fichiers DB3 exportés doivent alors être traités.
- Sauvegarder tous les fichiers exportés reçus en les archivant dans un sous-dossier du dossier C:\GATSAndroid2\. Il n'est pas nécessaire de les agréger à nouveau.

Durant et après l'agrégation, il importe de vérifier les points suivants :

- Tous les fichiers reçus présentent des noms de fichiers indiquant la date d'exportation du fichier, ce qui garantit que le traitement est effectué sur des fichiers en ordre. Si les fichiers sont reçus dans le désordre (c'est-à-dire si des fichiers de données antérieurs sont reçus ultérieurement), les données les plus récentes pour un numéro d'enquête particulier seront écrasées par des données antérieures potentiellement incomplètes. Afin de vous assurer, lors de l'agrégation vers un master.db3 non vide, que les données que vous avez traitées sont bien les plus récentes, reportez-vous aux fichiers DB3 archivés. (Agrégez toujours les fichiers par ordre chronologique, ou bien utilisez uniquement les fichiers les plus récents.)
- Exactitude et exhaustivité de l'agrégation des données : tous les fichiers DB3 escomptés ont-ils été pris en compte et inclus dans l'agrégation ?
- Pour chaque enquêteur de terrain, le fichier DB3 agrégé est-il bien le plus récent ?
- Examen des données : il convient de vérifier et de valider les données. L'analyste doit examiner les données d'événement ainsi que les questionnaires remplis afin de repérer les éventuels problèmes, les tendances à l'œuvre et les potentielles incohérences de données.
Cet examen des données permettra également d'établir des fréquences par question, pour examen. De plus amples informations au sujet de la visualisation des fichiers DB3 sont fournies dans le *Plan de mise en œuvre pour la gestion des données de l'enquête GATS*.

Examinez le rapport de synthèse de l'agrégation et le rapport d'état de l'agrégation générés lors de l'agrégation afin de confirmer les chiffres.

Le rapport de synthèse de l'agrégation des données fournit des informations sur les résultats de l'agrégation. Chaque ligne présente les résultats d'une unité d'agrégation (généralement un enquêteur de terrain) et d'une table de base de données spécifique. Le tableau est désigné dans la colonne AggTable, le nombre de lignes du tableau de saisie est indiqué dans la colonne AggRowsTotal, et le nombre de lignes insérées à partir de ce tableau est indiqué dans la dernière colonne (AggRowsInserted).

Le rapport d'état de la sélection des ménages est une synthèse de l'état de toutes les enquêtes que contient le fichier de données agrégées. Pour chaque unité d'agrégation (généralement un enquêteur de terrain), le nombre total d'enquêtes est rapporté pour cette unité dans la colonne Enquêtes. Le total est ensuite ventilé selon les différentes catégories d'état. Les colonnes relatives à l'état (Entretiens réalisés, Non traités, En attente et Finaux) dépendent du code de résultat de l'enquête le plus récent.

- En utilisant les données agrégées ou tout fichier DB3, les données au format fichier plat ou les fichiers CSV peuvent être créés à l'aide des fonctionnalités d'agrégation de données disponibles dans la Boîte à outils.
- Il est recommandé d'examiner les données d'événement (contenues dans les fichiers DU et DUEVT) pour acquérir une vue d'ensemble et une compréhension des opérations de terrain.
- Le fichier DU (unité de logement) contient seulement un enregistrement de chaque numéro d'enquête (le *questionnaire destiné aux ménages* et le *questionnaire individuel* ont chacun leur propre numéro d'enquête, à savoir -00 pour le premier et -01 pour le second) ainsi que le code de résultat le plus récent pour ce numéro d'enquête. Le fichier DUEVT contient l'historique des événements associés à chaque numéro d'enquête (renseigné par le ou les enquêteurs de terrain) et peut contenir plusieurs registres par numéro d'enquête.

Générer des fichiers de données pour examen

- Les données du questionnaire sont stockées dans une seule table (Table de réponses) réunissant à la fois les données du *questionnaire destiné aux ménages* et celles du *questionnaire individuel*. Cette table possède une ligne par question, donc les données doivent être réorganisées avant l'analyse. Les données des réponses doivent faire l'objet d'une transposition afin de pouvoir être facilement examinées et utilisées dans un logiciel statistique. Ce processus permet de créer un enregistrement fondé sur le numéro d'enquête pour les données des deux questionnaires (*questionnaire destiné aux ménages* et *questionnaire individuel*). Le logiciel GSS transpose les données des réponses à partir des fichiers DB3 de l'enquête spécifiés avec des fichiers séparés au niveau du numéro d'enquête et des fichiers plats, pour le *questionnaire destiné aux ménages* et le *questionnaire individuel*.
- Les données d'événement de l'enquête GATS sont conservées dans la table DUEVT, avec une ligne par résultat pour chaque numéro d'enquête. Les données relatives à l'unité de logement sont conservées dans la table DU, avec une ligne par numéro d'enquête. Les données d'unité de logement contiennent également une copie du code de résultat le plus récent pour chaque enquête.

Le tableau ci-dessous résume les différentes sources de données des dispositifs portables.

Source de données du tableau	Tableau	Ligne
Données du questionnaire	Réponses	Une ligne par question pour chaque numéro d'enquête et formulaire
Données d'événement	DUEVT	Une ligne par résultat pour chaque numéro d'enquête et formulaire
Unité de logement	DU	Une ligne par unité de logement
Modifications d'adresse	Historique d'adresses	Une ligne par modification d'adresse
Remarques de l'enquêteur de terrain	Remarques	Une ligne par type de remarque (par numéro d'enquête ou par question)

4.2.2 Contrôle qualité des données collectées

Au Centre national des données, des procédures de contrôle qualité doivent être en place afin d'examiner, de suivre, de rapporter et de sauvegarder régulièrement les données au fur et à mesure de leur arrivée et de leur traitement. Le Centre national des données doit établir un calendrier de traitement quotidien, au minimum à des fins d'établissement de rapports et de suivi, et qui prévoit également un processus régulier de sauvegarde.

Le protocole normalisé de l'enquête GATS comporte une liste de variables d'indicateurs de GATS et une suggestion de rapports au niveau national. De plus amples précisions au sujet des principales variables d'indicateurs sont fournies dans le document *Définitions des indicateurs de l'enquête GATS*. Les pays peuvent, le cas échéant, compléter/adapter les variables d'indicateurs dans le questionnaire final approuvé.

Il est recommandé d'examiner les variables des indicateurs clés au niveau des données brutes pendant la collecte des

Rapports élaborés par les pays : Il est recommandé que les pays établissent un mécanisme de rapports au sujet du statut des événements au niveau des enquêtes. Le suivi de l'enquête (progression sur le terrain, collecte et transmission des données, etc.) constitue une activité essentielle, coordonnée et continue.

- Durant la période de collecte de données, le Centre national des données doit fournir un rapport hebdomadaire sur le taux de réponse généré à partir du logiciel GSS. Les informations doivent préciser les entretiens réalisés et le taux de réponse global. Des exemples de rapports de terrain et de rapports d'évolution des taux de réponse sont fournis dans le *Guide de programmation sur General Survey System*.

- Le Centre national des données est encouragé à utiliser les rapports générés avec le logiciel GSS afin de rédiger des rapports dans Access^{1,2}, SPSS^{1,3}, SAS^{1,4}, STATA^{1,5} ou d'autres logiciels, selon les cas, dans le but de faciliter la détection d'incohérences, d'anomalies ou de lacunes lors de l'examen des données. Les données agrégées transposées doivent être utilisées pour générer des rapports.

Les incohérences dans les données peuvent être identifiées en effectuant des vérifications de validation entre les variables au sein du *questionnaire destiné aux ménages* et du *questionnaire individuel* et entre eux. Pour en savoir plus sur les vérifications de validation, veuillez vous référer au *Chapitre 5* consacré aux mesures d'assurance qualité à réaliser à l'issue de la collecte de données.

Procédures de sauvegarde : Les données doivent être sauvegardées de manière régulière, par exemple à échéance hebdomadaire, selon la fréquence d'agrégation.

- Les fichiers DB3 individuels (au niveau des enquêteurs de terrain) doivent être archivés.
- Les données archivées doivent également être sauvegardées sur un support secondaire, tel qu'un réseau ou une clé USB, dont une copie doit être conservée en deux endroits physiques distincts.
- Une fois la collecte de données terminée, l'harmonisation des données reçues doit s'effectuer à partir du fichier de sélection de l'échantillon-maître ainsi que des fichiers exportés du terrain.

4.2.3 Communication et appui technique

Durant la collecte des données, le Centre national des données doit communiquer des rapports d'étape hebdomadaires à l'OMS et au CDC et signaler tout problème technique. Tous les problèmes techniques liés aux dispositifs portables ou au logiciel doivent être signalés par courrier électronique le plus tôt possible, même s'ils ont été résolus. Cette démarche contribuera à faciliter l'assistance technique, la résolution des problèmes et le partage d'informations en temps utile.

¹ L'utilisation de noms commerciaux sert uniquement à des fins d'identification et n'implique aucunement l'approbation du département de la Santé et des Services sociaux des États-Unis d'Amérique ou des Centers for Disease Control and Prevention.

² Microsoft Office Access® (Microsoft Corporation, Redmond, Washington).

³ SPSS® (SPSS Inc., Chicago, Illinois).

⁴ SAS® (SAS Institute Inc., Cary, Caroline du Nord).

⁵ STATA® (Stata Corp., College Station, Texas).

5. Assurance qualité : Actions à mener après la collecte des données

La phase postérieure à la collecte des données comprend une série de tâches à accomplir afin de préparer un fichier de données analytiques en vue de l'analyse des données. Cette phase intervient une fois que toutes les données de l'enquête ont été collectées et agrégées. Cette phase recouvre i) la préparation des données pour le calcul des pondérations d'échantillon, ii) l'évaluation de la qualité de l'échantillonnage, de l'erreur d'échantillonnage et des pondérations, et iii) la mesure de la qualité de la couverture de la population de l'enquête, des taux de non-réponse et d'autres erreurs non liées à l'échantillonnage. Cela inclut également la préparation d'un fichier de données analytiques. Ce chapitre décrit les directives et les procédures d'assurance qualité applicables et recommandées pour chacune de ces activités, en abordant spécifiquement les points suivants :

- 5.1 Nettoyage et préparation des données pour le calcul des pondérations de l'échantillon
 - 5.1.1 Création de la base de données principale
 - 5.1.2 Élimination des variables d'identification confidentielles
 - 5.1.3 Nettoyage et validation du fichier de données fusionné
 - 5.1.4 Attribution des codes de décision finale
- 5.2 Mesures de qualité : échantillonnage, erreur d'échantillonnage et pondération
 - 5.2.1 Tendance des ajustements d'étalonnage des pondérations par stratification a posteriori entre les cellules d'ajustement
 - 5.2.2 Effet multiplicateur des variables de pondération de l'échantillon sur la précision des estimations de l'enquête
 - 5.2.3 Effet du plan de sondage global sur la précision des estimations de l'enquête et au niveau intra-UPE
 - Homogénéité intra-groupe des principales estimations correspondantes de l'enquête
 - 5.2.4 Marge d'erreur pour les principales estimations de l'enquête
- 5.3 Mesures de la qualité : couverture, non-réponse et autres erreurs non liées à l'échantillonnage
 - 5.3.1 Taux de couverture de la base de ménages
 - 5.3.2 Tendances de taux d'abandon des répondants
 - 5.3.3 Tendances des taux de réponse des ménages par strate d'échantillonnage de la première phase
 - 5.3.4 Tendances des taux de réponse au niveau individuel parmi les variables utilisées pour les ajustements pour non-réponse
 - 5.3.5 Tendances des taux de refus au niveau individuel parmi les variables utilisées pour les ajustements pour non-réponse
 - 5.3.6 Taux de non-réponse par question pour les variables d'indicateurs de la fiche d'information
- 5.4 Examen formel de la qualité statistique
 - 5.4.1 Documentation de la pondération de l'échantillon
- 5.5 Création du fichier de données analytiques
- 5.6 Création du fichier de données destiné à une utilisation publique

5.1 Nettoyage et préparation des données pour le calcul des pondérations de l'échantillon

Cette section fournit des directives au sujet de la fusion des fichiers de données et de la validation des variables et des instructions « passer à », ainsi que pour la création des codes de décision à l'aide d'un logiciel statistique¹.

5.1.1 Création de la base de données principale

L'option « Création des fichiers d'analyse » rassemble les données d'enquête dans un seul fichier délimité par des virgules et génère des programmes de saisie SAS, SPSS et STATA capables de lire la base de données principale nouvellement fusionnée. Ces programmes peuvent être créés à l'aide du logiciel GSS.

Les fichiers utilisés dans l'option « Générer des fichiers d'analyse » comprennent le fichier maître DB3 agrégé final créé par agrégation des fichiers DB3 de tous les enquêteurs, ainsi que les deux bases de données de questionnaires (Survey0.db3 et Survey1.db3).

Lorsque vous générez des fichiers d'analyse, assurez-vous d'utiliser les emplacements les plus récents du fichier agrégé ainsi que des bases de données du questionnaire de l'enquête. Pour modifier ce paramètre, vous pouvez mettre à jour l'emplacement ou le nom du fichier dans la zone de texte, ou bien utiliser le bouton « ... » et sélectionner un fichier existant parmi ceux proposés par le système. La base de données principale sera un fichier délimité par des virgules au format Unicode (UTF-8). L'option générera également un code de saisie pour la lecture des données dans SAS, SPSS et STATA.

Pour visualiser le fichier CSV, il est nécessaire d'utiliser un éditeur de texte Unicode, par exemple WordPad. Les fichiers SAS, SPSS et STATA sont également en Unicode afin de s'assurer que les étiquettes et formats s'affichent correctement. Il peut s'avérer nécessaire de mettre à jour les programmes SAS, SPSS et STATA avant de les utiliser, afin de s'assurer qu'ils sont bien configurés selon les paramètres spécifiques au pays. Ce processus permet la création d'un ensemble de données principal. Cette base de données principale créée doit être utilisée pour retirer toute variable d'identification confidentielle et doit être partagée avec le référent informatique dans le pays.

5.1.2 Élimination des variables d'identification confidentielles

Sur l'ensemble de données principal ainsi créé, supprimez les champs d'identification confidentiels des champs du fichier de sélection de l'échantillon-maître.

5.1.3 Nettoyage et validation du fichier de données fusionné

Vérifiez que les variables ont des valeurs appropriées, vérifiez que les instructions « passer à » fonctionnent correctement et vérifiez les champs vides. Bien qu'un grand nombre de ces contrôles de qualité des données soient mis en œuvre au moment de l'administration de l'entretien par le logiciel portable, il est important de les vérifier à nouveau au cas où des erreurs non détectées auraient été commises dans la programmation du logiciel portable.

¹Le Centre de coordination des données de l'enquête GATS au sein du CDC fournit une assistance technique pour l'utilisation des logiciels de statistiques suivants : SAS, SPSS et STATA.

1. Vérifiez que les instructions « passer à » fonctionnent tel que spécifié dans le questionnaire. Les instructions « passer à » originales dérivées du questionnaire national final approuvé doivent être vérifiées pour toutes les variables principales ; il est également recommandé de vérifier les instructions pour toutes les autres variables.

- Un répondant qui n'a pas répondu à une question en raison d'une instruction « passer à » aura un champ vide pour cette question restée sans réponse.
 - Par exemple, si un répondant n'a pas souhaité répondre à la question A02a (le mois de naissance du répondant) ou A02b (l'année de naissance du répondant), il doit répondre à la question A03 (l'âge du répondant), mais si des réponses sont apportées aux questions A02a et A02b, alors il faudra ignorer la question A03.
 - Par exemple, à la question B01, si un répondant répond fumer « quotidiennement » (valeur 1), il passera alors à la question B04 et n'aura pas besoin de répondre aux questions B02 et B03.
- Dans SAS, un champ vide pour une variable numérique s'affiche « . » (pour les variables de caractères, un champ vide s'affiche « »). Dans SAS, l'instruction « passer à » peut être vérifiée en utilisant le code ci-dessous. Tous les registres qui ne respectent pas les instructions « passer à » sont exportés vers un ensemble de données SAS.

```
SI B01 = 1 & (B02 NE . OU B03 NE .) ALORS AFFICHER=1 ;  
SINON SI B01 = 2 & (B02 = . OU B03 = .) ALORS AFFICHER =1 ;  
SINON SI B01 = 3 & (B02 NE . OU B03 = .) ALORS AFFICHER =1 ;  
SINON AFFICHER=0 ;  
ALORS AFFICHER=1 PUIS EXPORTER ; *La valeur 1 signifie que l'instruction « passer  
à » n'a pas fonctionné et la valeur 0 signifie que l'instruction a fonctionné ;
```

2. Vérifiez que tous les champs vides résultent d'une instruction « passer à » (ce qui signifie que ces champs ne concernent pas le répondant et doivent effectivement rester vides). Tout autre champ vide est non valide.
 - Une nouvelle valeur devra être attribuée aux champs laissés vides en raison d'une instruction « passer à », afin de les distinguer des champs vides qui ne sont pas attribuables à des instructions « passer à » (champs vides non valides). Comme les valeurs des variables originales ne doivent pas être écrasées, créez un nouvel ensemble de données pour réaliser cette étape.
 - Dans SAS, utilisez le code « .S » pour signaler une valeur manquante due à une question non applicable (champ vide valide). Le code SAS pour illustrer l'instruction « passer à » utilisée ci-dessus peut alors être modifié comme suit :

```
SI B01 = 1 & (B02 = . OU B03 = .) ALORS B02 = .S & B03 = .S ;  
SINON SI B01 = 3 & B02 = . ALORS B02 = .S ;
```
 - Une fois que toutes les valeurs des instructions « passer à » ont été remplacées, il ne devrait plus rester de champs vides.
S'il en reste, il faut alors exporter ces registres vers le fichier d'erreur.

3. Vérifiez chaque variable afin de vous assurer qu'il n'existe aucune valeur non valide. Utilisez le dictionnaire des données du pays (fichier de codes) afin de connaître les valeurs valables pour chaque variable de l'ensemble de données.
 - Vérifiez que les valeurs des réponses à chaque question sont valides.
 - Par exemple : Les seules réponses valides à la question B01 sont : 1 (« tous les jours »), 2 (« à une fréquence moindre »), 3 (« jamais »), -9 (« n'a pas souhaité répondre »), ou le code de valeur manquant spécifique au logiciel utilisé (par exemple, .S sur SAS).
 - Par exemple : Certaines variables sont associées à des plages, de sorte qu'aucune valeur inférieure ou supérieure à la plage indiquée ne peut être saisie. Ainsi, la variable B04 est associée à une plage comprise entre 0 (valeur minimale) et 98 (valeur maximale). Une valeur de -9 est également admise pour les questions auxquelles il a été répondu « Ne sait pas » et « N'a pas souhaité répondre ». Le code de valeur manquant spécifique au logiciel utilisé (par exemple, .S sur SAS) serait également une valeur valide.
 - *Dans SAS, les registres comportant des réponses non valides à la question B01 peuvent être sélectionnés et exportés vers le fichier d'erreurs en utilisant le code suivant :*

```
SI B01 N'EST PAS COMPRIS DANS LE CHAMP (1,2,3,-7,-9) ALORS
EXPORTER ;
```
 - *Si l'analyste le souhaite, il est aussi possible d'appliquer des fréquences simples à chacune des variables afin de vérifier qu'aucune valeur n'est en dehors de la plage définie. Tous les registres présentant des valeurs non valides peuvent alors être identifiés et exportés vers le fichier d'erreurs.*
 - Vérifiez l'âge du répondant.
 - Afin de vérifier que le répondant a l'âge approprié, il faut calculer son âge et l'intégrer dans l'ensemble de données. Pour cela, on calcule l'écart entre la date de l'entretien individuel (IQ_EVENTDATE) et le mois (A02a) et l'année (A02b) de naissance des répondants, le jour étant supposé être le 15 du mois. Si un répondant a indiqué « Ne sait pas » ou « N'a pas souhaité répondre » aux questions A02a ou A02b, il lui est alors demandé d'indiquer son âge à la question A03 (la question A03 ne s'applique qu'en l'absence de réponse aux questions A02a ou A02b). Si le mois ou l'année de naissance ne sont pas indiqués, utilisez la question A03 pour indiquer l'âge du répondant. Si le mois et l'année sont indiqués, calculez l'âge en soustrayant la date de naissance de la date de l'entretien. Vérifiez que l'âge du répondant est valide (valeur comprise en 15 et 125).
 - *Dans SAS, l'âge peut être calculé puis vérifié en utilisant le code suivant :*

```
SI A02A ≤ 12 ET A02B ≠ -7 ET A02B ≠ -9 ALORS APPLIQUER ;
BIRTHDATE = MJA(A02A, 15, A02B) ;
AGE = FLOOR(YRDIF(BIRTHDATE, IQ_EVENTDATE, 'ACTUAL')) ;
TERMINER ;
SINON AGE = A03 ;
```

4. Autres vérifications de validation des données :

- Si l'une des variables suivantes a été appliquée, vérifiez que sa valeur est inférieure à l'âge du répondant : B04, B05, B08, B09, B11, B12, B13, C04, C05, C08, C09, C11, C12 et C13 (selon l'unité de mesure employée, il peut s'avérer nécessaire de préalablement convertir ces valeurs en années).

5.1.4 Attribution des codes de décision finale

Le fichier de données fusionné comprendra les codes de résultats définitifs qui ont été attribués par les enquêteurs de terrain à toutes les enquêtes de l'échantillon réalisées durant la collecte des données (se reporter au *Manuel GATS d'enquête sur le terrain* pour la description des codes de résultat sur le terrain). Chaque entretien auprès d'un ménage et chaque entretien individuel doit avoir un code de résultat définitif (voir la *section 4.1.3* au sujet de l'attribution des codes de résultats définitifs pour les ménages non interrogés durant la collecte des données).

Des codes de décision finale doivent être attribués à chaque dossier. Les codes de décision finale sont utilisés pour calculer et rapporter les taux de réponse ainsi que les mesures

Les codes de décision finale doivent être attribués à partir des codes de résultats définitifs. Les codes de décision finale sont ensuite utilisés pour calculer et rapporter les taux de réponse ainsi que les mesures d'assurance qualité. Reportez-vous à l'*annexe C.1* pour l'attribution des codes de décision finale à partir des codes de résultats définitifs et pour obtenir les calculs des taux de réponse ainsi qu'un modèle pour rapporter ces résultats de la collecte de données.

Voici quelques directives importantes à suivre dans le cadre de ce processus :

1. À chaque enquête doit être attribué un code de décision finale pour le *questionnaire destiné aux ménages*, à partir du code de résultat définitif sur le terrain pour ce questionnaire. Si l'entretien associé au *questionnaire destiné aux ménages* a été réalisé et qu'une personne a été sélectionnée pour l'entretien individuel (code de décision finale égal à 1), un code de décision finale doit également être attribué pour le *questionnaire individuel*. Si le code de décision finale du *questionnaire destiné aux ménages* est différent de 1 (par exemple, si aucune personne n'a été sélectionnée, si le ménage n'a pas souhaité répondre ou si le questionnaire est incomplet), il n'y aura pas de code de décision finale pour le *questionnaire individuel*, étant donné qu'aucun entretien individuel n'a été effectué.
2. Le *questionnaire individuel* comporte certaines questions essentielles permettant de déterminer la validité de l'entretien. Si l'une de ces questions essentielles a reçu pour réponse « Ne sait pas » ou « N'a pas souhaité répondre », alors l'entretien n'est pas considéré comme valable. Ainsi, pour toutes les enquêtes individuelles dont le code de résultat définitif est 400 (*questionnaire individuel* rempli), il convient d'appliquer les règles ci-dessous pour déterminer le code de décision à attribuer.
 - Si les questions B01, B02 ou B03 (questions essentielles sur la prévalence du tabagisme) ont reçu pour réponse « Ne sait pas » ou « N'a pas souhaité répondre », un code de décision finale de 16 (Répondant sélectionné non qualifié) doit être attribué. Cependant, ceci ne doit pas être une pratique courante. (Cette règle peut également être appliquée pour

- C01, C02 et C03 (questions essentielles sur la prévalence du tabac sans fumée), uniquement si la consommation de tabac sans fumée est courante dans le pays).
- Pour toutes les autres enquêtes individuelles dont le code de résultat définitif est de 400 (*questionnaire individuel* rempli), il convient d'attribuer le code de décision finale de 11 (*questionnaire individuel* rempli). La plupart des enquêtes répondront à ce cas de figure.
3. Pour toutes les enquêtes individuelles dont le code de résultat définitif est 402 (*questionnaire individuel* partiellement rempli), il convient d'appliquer les règles ci-dessous pour déterminer le code de décision à attribuer.
- Si le *questionnaire individuel* a été rempli au moins jusqu'à la question E01 et qu'aucune des questions essentielles (énoncées au point 2 ci-dessus) n'a donné lieu à la réponse « Ne sait pas » ou « N'a pas souhaité répondre », un code de décision finale de 11 (*questionnaire individuel* rempli) doit être attribué.
 - Si le *questionnaire individuel* a été interrompu avant la question E01, un code de décision finale de 12 (Questionnaire incomplet) doit être attribué. (Les enquêtes auxquelles est attribué un code de décision finale de 12 [Questionnaire incomplet] sont considérées comme des non-réponses et les données de ces répondants ne sont pas prises en compte dans l'analyse).
 - Si le *questionnaire individuel* a été rempli au moins jusqu'à la question E01 mais que la réponse donnée à l'une des questions essentielles (énoncées au point 2 ci-dessus) est « Ne sait pas » ou « N'a pas souhaité répondre », un code de décision finale de 16 (Répondant sélectionné non qualifié) doit être attribué. (Comme expliqué au point 2, les répondants qui ne sont pas en mesure de fournir de réponse valide aux questions essentielles ne sont pas considérés comme qualifiés pour participer à l'enquête. Par conséquent, les données de leur entretien ne sont pas prises en compte dans l'analyse).
4. Seules les enquêtes présentant un code de décision de 11 doivent être intégrées dans l'ensemble de données analytiques final (d'où l'importance d'attribuer correctement les codes de décision).
5. Utilisez les tabulations croisées pour vérifier tous les codes de résultats définitifs par rapport aux codes de décision, afin de détecter les erreurs de classification.
L'absence de correspondance entre les deux codes témoigne d'un problème avec le code du logiciel utilisé pour créer les codes de décision.

5.2 Mesures de la qualité : échantillonnage, erreur d'échantillonnage et pondération

Cette section décrit les calculs permettant d'évaluer directement la qualité des estimations provenant des échantillons GATS et d'indiquer les effets des grappes d'échantillonnage ainsi que des inégalités de pondération sur ces estimations. Des directives concernant la vérification de l'exactitude des pondérations calculées y sont également fournies.

5.2.1 Tendances des ajustements d'étalonnage des pondérations par stratification a posteriori entre les cellules d'ajustement

La dernière étape de production des pondérations de l'échantillon consiste à calibrer les pondérations sur les effectifs de la population, en fonction de corrélats connus des principales mesures de résultats de l'étude, appelés variables d'étalonnage (par exemple, le sexe, le niveau d'études, l'âge, le milieu urbain/rural et la région, comme le suggère le *Manuel de pondération de l'échantillon de l'enquête GATS*).

Des informations contextuelles ainsi que des instructions pour le calcul des ajustements après stratification sont fournies plus loin dans cette section. De plus amples détails sur ces sujets sont fournis à l'*annexe C.2*.

1. **Contexte** □ L'étalonnage permet de corriger le déséquilibre de l'échantillon non compensé par l'ajustement pour non-réponse. Des ajustements distincts sont appliqués à tous les membres des cellules d'ajustement créées stratégiquement. On souhaite augmenter les pondérations des sous-catégories de population qui demeurent sous-représentées et réduire celles des sous-catégories encore surreprésentées. Plus les valeurs des ajustements d'étalonnage s'écartent de 1,00 (que ce soit au-delà ou en deçà), plus l'impact potentiel du déséquilibre de l'échantillon (au-delà de celui pris en compte par l'ajustement pour non-réponse) sur le biais des estimations de l'enquête sera important.
2. **Réalisation d'ajustements par stratification a posteriori** □ Selon la procédure, l'étalonnage par stratification a posteriori implique la création de « cellules d'ajustement » par la classification croisée des mesures corrélées. L'ajustement par stratification a posteriori dans chacune de ces cellules d'ajustement est inférieur à 1 si le sous-échantillon de cette catégorie était surreprésenté dans l'échantillon après prise en compte de la sélection et du taux de non-réponse de l'échantillon, et supérieur à 1 dans les catégories où l'échantillon était sous-représenté.
3. **Rapport des ajustements par stratification a posteriori** □ Créez un tableau qui répertorie toutes les cellules d'ajustement, en précisant pour chacune d'elles la manière dont elles sont définies en fonction des variables catégorielles utilisées pour l'étalonnage. Pour chaque cellule, relevez la valeur de l'ajustement par stratification a posteriori calculée et observez sa taille par rapport à 1,00. L'idéal est que tous les ajustements par stratification a posteriori soient proches de 1, certains étant légèrement supérieurs à 1 et les autres légèrement inférieurs à 1,00.

5.2.2 Effet multiplicateur des variables de pondération de l'échantillon sur la précision des estimations de l'enquête

Le *Manuel d'échantillonnage de l'enquête GATS* préconise un plan d'échantillonnage où les probabilités de sélection (et donc la pondération de l'échantillon) varieront légèrement en raison du recours à des estimations de la taille des grappes, d'ajustements des pondérations d'échantillon, ainsi que de la répartition équitable des tailles d'échantillon entre les régions, lorsque des estimations régionales respectant les critères de précision de l'enquête GATS sont nécessaires. Le *Manuel GATS de pondération de l'échantillon* décrit la manière dont ces pondérations doivent être calculées. Une fois que les données du questionnaire ont été nettoyées et que les pondérations finales de l'échantillon ont été jointes, les données de l'échantillon GATS sont alors prêtes pour l'analyse et pour le calcul de cet effet.

Des informations contextuelles ainsi que des instructions pour le calcul de cet effet sont fournies plus loin dans cette section. De plus amples détails sur ces sujets sont fournis à l'*annexe C.3*.

1. **Contexte** □ La variabilité des pondérations de l'échantillon peut augmenter la quantité d'erreurs d'échantillonnage dans les estimations de l'enquête, entraînant des variances et des erreurs-types plus importantes pour ces estimations. Plus précisément, l'augmentation multiplicatrice de la variance des estimations de l'enquête, appelée $Meff_w$, dépend de la variabilité des pondérations de l'ensemble des observations de l'échantillon qui sont mises à profit afin de produire l'estimation.

Plus les pondérations sont variables, plus la valeur de $Meff_w$ sera élevée.
2. **Estimation de $Meff_w$** □ La moyenne simple et la variance de la pondération pour l'ensemble des observations de l'échantillon mises à profit pour produire les estimations de l'enquête sont nécessaires pour le calcul de $Meff_w$.
 - La valeur de $Meff_w$ pour les estimations tirées de cet ensemble d'observations de l'échantillon est mesurée en calculant d'abord le rapport de la variance sur le carré de la moyenne, puis en ajoutant un à ce rapport. La formule de calcul spécifique est l'Équation (1) à l'*annexe C.3*.
3. **Rapport concernant $Meff_w$** □ Puisque $Meff_w$ s'applique à toutes les estimations tirées de l'ensemble spécifique d'observations de l'échantillon définissant un domaine de rapport pour les estimations de l'enquête (par exemple, la valeur globale $Meff_w$, la population entière, la population rurale, le niveau d'études, l'âge, le genre, etc.),
 - la formule doit être calculée et rapportée pour tous les sous-groupes principaux de population pour lesquels des estimations de l'enquête GATS seront calculées et rapportées. Cela peut être effectué dans un tableau avec une liste de rapports des sous-groupes et les valeurs associées de $Meff_w$.
4. **Interprétation de $Meff_w$** □ Une interprétation correcte de la valeur de $Meff_w$ calculée pour un domaine de rapports est la suivante :

Cela est préférable pour $Meff_w < 2.00$.

 - « La variation de la pondération de l'échantillon a augmenté la variance de toutes les estimations (du domaine de rapports) par un facteur de ($Meff_w$). »
 - Si $Meff_w < 2.00$, les efforts pour réduire la variation de pondération sont inutiles.
5. **Exemple** □ Supposons que pour les estimations en milieu rural $Meff_w = 1.62$

« La variation de la pondération de l'échantillon a augmenté la variance de toutes les estimations des répondants ruraux par un facteur de 1,62. »

 - $Meff_w = 1.62$ est considéré comme suffisamment faible pour ces estimations.

5.2.3 Effet du plan de sondage global sur la précision des estimations de l'enquête et sur l'homogénéité intragroupe dans les unités primaires d'échantillonnage (UPE) des principales estimations correspondantes de l'enquête

Le *Manuel d'échantillonnage de l'enquête GATS* recommande une conception où les grappes des répondants à l'enquête sont sélectionnées ; le *Manuel de pondération de l'échantillon de l'enquête GATS* décrit comment calculer la pondération de l'échantillon. Une fois que les données du questionnaire ont été nettoyées et que les pondérations finales de l'échantillon ont été calculées, les données de l'échantillon GATS sont alors prêtes à être examinées avant le rapport des conclusions de l'analyse.

Le contexte général et les instructions pour le calcul de l'effet du plan de sondage global sur une estimation de l'enquête et de l'homogénéité au sein des UPE sont présentés plus loin dans cette section. De plus amples détails sur ces sujets sont fournis à l'*annexe C.4*.

1. **Contexte** — La variance d'une estimation de l'enquête ($\hat{\theta}$) d'un plan d'échantillonnage complexe, divisée par la variance d'une estimation comparable basée sur un échantillon aléatoire simple de même taille est appelée l'« effet du plan de sondage » global de $\hat{\theta}$, ou $Deff(\hat{\theta})$. Il existe deux composantes multiplicatrices de $Deff(\hat{\theta})$: l'effet de plan de sondage dû à l'échantillonnage en grappes, $Deff_{CS}(\hat{\theta})$, et l'effet multiplicateur de la variable de pondération de l'échantillon, $Meff_w$ (voir la section 5.2.2). Par ailleurs, $Deff_{CS}(\hat{\theta})$ est directement liée au taux d'homogénéité au sein des UPE, ρ_{θ} , et à la taille moyenne des grappes de l'échantillon.
2. **Estimation de $Deff(\hat{\theta})$ et ρ_{θ}** — Une estimation de $Deff(\hat{\theta})$ peut être obtenue directement avec des ensembles de logiciels d'analyse d'enquête (par exemple, SUDAAN), ou peut être calculée à partir de l'estimation et de sa variance lorsque l'estimation est un rapport ou un taux. Puisque $Deff(\hat{\theta})$ dépend de $Deff_{CS}(\hat{\theta})$, ρ_{θ} , et de la taille moyenne des grappes de l'échantillon, on peut obtenir une estimation de ρ_{θ} en la résolvant et en utilisant les estimations de $Deff(\hat{\theta})$ ainsi que ces autres mesures.
 - Des formules détaillées pour estimer $Deff(\hat{\theta})$ (lorsque le logiciel d'analyse de l'enquête ne remplit pas cette fonction) et pour estimer ρ_{θ} en toutes circonstances sont données dans les Équations (1) et (2) à l'*annexe C.4*. Les estimations obtenues sont notées $deff(\hat{\theta})$ et $\hat{\rho}_{\theta}$, respectivement. On utilise $Deff_{CS}(\hat{\theta})$ pour désigner l'estimation de $deff_{CS}(\hat{\theta})$.
 - Puisqu'il y aura de nombreuses valeurs de $deff(\hat{\theta})$ et $\hat{\rho}_{\theta}$, synthétisez-les en rapportant les valeurs médianes, minimales et maximales.
3. **Rapports de $Deff(\hat{\theta})$ et de ρ_{θ}** — Les estimations de $Deff(\hat{\theta})$ et de ρ_{θ} doivent être rapportées pour toutes les principales mesures de l'étude (voir l'*annexe C.4* pour les définitions).
4. **Interprétation de $Deff(\hat{\theta})$ et de ρ_{θ}** — Une interprétation correcte des valeurs estimées de $deff(\hat{\theta})$ et ρ_{θ} est la suivante :

- « Compte tenu du plan d'échantillonnage de l'enquête GATS, la variance de l'estimation de l'enquête (de la caractéristique de la population θ) est $deff(\hat{\theta})$ fois supérieure à la variance qui aurait été obtenue si un échantillonnage aléatoire simple avait produit le même nombre de répondants. ».
 - Il est préférable que l'effet du plan de sondage global pour une estimation, ou $Deff(\hat{\theta})$, soit inférieur à 2,50.
 - La somme de la variation totale de la mesure de l'étude (associée à θ) qui n'a pas été comptabilisée dans les UPE est $\hat{\rho}_\theta$.
 - Il est rare que $\hat{\rho}_\theta > 0.15$ ou négatif ; généralement, $\hat{\rho}_\theta < 0.05$ voire $\hat{\rho}_\theta < 0.01$.
5. **Exemple** — L'analyse pour estimer l'actuel taux de prévalence du tabagisme, p , produit les résultats suivants à partir d'un échantillon dont la taille moyenne des grappes de l'échantillon était de 20 : $deff(\hat{\theta}) = 1.92$, $Meff_w = 1.36$.

Il est rare que
 $\hat{\rho}_\theta > 0.15$;
généralement,
 $\hat{\rho}_\theta < 0.05$ voire

« Compte tenu du plan d'échantillonnage de l'enquête GATS, la variance de l'estimation de l'enquête concernant l'actuel taux de prévalence du tabagisme est 1,92 fois supérieure à la variance qui aurait été obtenue si un échantillonnage aléatoire simple avait produit le même nombre de répondants. »

- La valeur calculée de $\hat{\rho}_\theta = 0.022$ est représentative.

5.2.4 Marge d'erreur pour les principales estimations de l'enquête

Il existe plusieurs manières de rendre compte de la précision statistique des estimations de l'enquête. L'une d'elles est la marge d'erreur d'une estimation. Le GATS conseille de rapporter la marge d'erreur estimée (MOE) ainsi que des estimations des principales mesures de l'enquête. Le *Manuel d'échantillonnage de l'enquête GATS* décrit les trois principales caractéristiques des échantillons du GATS qui vont influencer la qualité statistique des estimations et des conclusions tirées des données. Ces caractéristiques sont le recours à la stratification, la sélection de grappes de membres de la population cible, et la sélection de membres de la population avec des probabilités inégales (nécessitant donc l'utilisation de facteurs de pondération dans l'analyse). Divers logiciels d'analyse d'enquête peuvent être utilisés pour tenir compte de ces caractéristiques dans l'analyse des données de l'enquête GATS.

Des informations générales et des instructions de calcul sont présentées dans cette section. De plus amples détails sur ces sujets sont fournis à l'*annexe C.5*.

1. **La marge d'erreur est spécifique à la mesure** — Chaque estimation obtenue à partir des données de l'enquête GATS a sa propre marge d'erreur. La marge d'erreur est par définition la moitié de la largeur attendue d'un intervalle de confiance d'une estimation d'une principale mesure de l'enquête. La marge d'erreur peut également être interprétée comme l'écart entre l'estimation réalisée et la valeur réelle au sein de la population.
2. **Principales mesures de l'enquête** — Les estimations de changement d'une phase à une autre ainsi que les estimations spécifiques des phases de l'enquête GATS seront calculées pour divers types de comportements vis-à-vis du tabac et divers niveaux de consommation de tabac. Ces

estimations sont susceptibles d'être rapportées pour la population dans son ensemble, ou bien suivant différents sous-groupes étudiés, par exemple par sexe, milieu urbain/rural, et catégories d'âge. Des estimations régionales peuvent être rapportées pour les échantillons régionaux dont l'effectif global excède 8 000 individus.

Les trois indicateurs suivants ont été considérés comme les plus importants pour la planification :

- *Le taux de prévalence du tabagisme* □ le pourcentage de personnes qui consomment actuellement des produits du tabac ;
- *Le taux de tabagisme passif* □ le pourcentage de personnes qui sont exposées à la fumée produite par d'autres fumeurs ;
- *Le taux de sevrage* □ le pourcentage de fumeurs actuels qui ont essayé d'arrêter de fumer.

3. **Estimation de la marge d'erreur** — Le logiciel d'analyse de l'enquête ne calcule généralement pas d'estimation de la marge d'erreur, mais les informations nécessaires pour la calculer sont généralement disponibles. Trois éléments sont nécessaires pour calculer et interpréter la : i) l'estimation de la mesure de l'enquête, ii) l'estimation de son erreur-type (ou variance), et iii) une mesure spécifique associée au niveau de confiance statistique souhaité pour la valeur de la marge d'erreur.

Le niveau de confiance est généralement basé sur une valeur (Z) de la distribution normale standard.

Par exemple, pour un niveau de confiance de 95 %, nous pouvons utiliser $Z = 1,96$.

La marge d'erreur correspond à la mesure du niveau de confiance désirée multipliée par l'erreur-type de l'estimation. La formule réelle de calcul est l'Équation (1) fournie à l'*annexe C.5*.

4. **Rapports MOE** — Il est recommandé de calculer la MOE pour toutes les estimations liées aux principales mesures de l'enquête. Cela inclut les estimations nationales globales de ces mesures, ainsi que les estimations de ces mesures pour tous les principaux sous-groupes de rapport (par exemple selon le sexe, le milieu urbain/rural, la région et l'âge).

**Une marge d'erreur
approximative doit être calculée
pour toutes les estimations
associées aux principales
mesures de l'enquête.**

- Les principales estimations de l'enquête et leurs valeurs associées de MOE doivent être présentées ensemble. Des informations sur les rapports des estimations des erreurs d'échantillonnage et un exemple de matrices de tableau avec des indicateurs clés de l'enquête GATS et des valeurs recommandées pour les rapports sont fournis à l'*annexe C.6*.

5. **Interprétation** — Lorsqu'elle est analysée aux côtés de la valeur d'une estimation de l'enquête, MOE indique à quel point l'estimation est susceptible d'être proche de la valeur réelle au sein de la population.

Si l'on utilise $Z=1,96$ pour calculer MOE, l'estimation de l'enquête et la valeur de sa marge d'erreur peuvent être interprétées comme suit :

« Nous sommes sûrs à 95 % que l'estimation (VALEUR DE L'ESTIMATION) se trouve à (VALEUR DE LA MOE ASSOCIÉE) de la valeur au sein de la population correspondante. »

6. **Exemple** — Supposons que la valeur rapportée d'une estimation de l'enquête GATS soit de 43,4 %, avec une erreur-type de 1,5611 % qui a été calculée conformément au plan d'échantillonnage réel de ce pays.

La valeur de MOE pour cette estimation sera de $3.1\% = (1.96)(1.5611\%)$

Interprétation :

« Nous sommes sûrs à 95 % que l'estimation 43,4 % se trouve dans une marge de 3,1 % de la valeur au sein de la population correspondante. »

5.3 Mesures de la qualité : couverture, non-réponse et autres erreurs non liées à l'échantillonnage

5.3.1 Taux de couverture de la base de ménages

Comme l'indique le *Manuel d'échantillonnage de l'enquête GATS*, l'avant-dernière étape de l'échantillonnage de l'enquête GATS consiste à sélectionner un échantillon de ménages parmi ceux figurant dans les UPE ou les segments. La qualité de l'échantillon de ménages et, en fin de compte, des estimations de l'enquête GATS, dépendent donc de l'exhaustivité de ces listes de ménages, lesquelles sont dressées de diverses manières.

Des informations contextuelles ainsi que des instructions pour le calcul du taux de couverture de la base de ménages sont présentées plus loin dans cette section. De plus amples détails sur ces sujets sont fournis à l'*annexe C.7*.

1. **Contexte** □ Une étape cruciale du processus d'échantillonnage d'une enquête auprès des ménages est le recensement des ménages au cours de la phase d'échantillonnage, lors du choix des adresses des ménages. Différentes approches de recensement des ménages peuvent être utilisées (constitution classique d'une liste des ménages par des agents recenseurs sur le terrain, listes électorales modifiées, etc.). Le résultat final habituel de telles approches est un sous-recensement net des ménages et donc une couverture incomplète dans l'échantillon final des ménages. Puisque l'effet de biais de la sous-couverture dépend de i) la différence cumulée entre les principales mesures de l'étude pour les ménages couverts par la base de sondage et ceux non couverts et ii) l'ampleur de la non-couverture (ou 100 % moins le taux de couverture des ménages) dans l'échantillon, il est utile d'estimer le taux de couverture de la base de sondage pour les ménages.
 - Le taux de couverture d'une base de sondage (FCR en anglais) est défini comme le nombre de ménages pris en compte dans la base de sondage des ménages divisé par le nombre de ménages dans la population.
 - S'il est utile de disposer de données pour décrire la couverture de la base de sondage de l'échantillon des ménages, il est rare que l'on dispose de données provenant d'un recensement récent du nombre total de ménages au sein de la population. Si de telles données sont disponibles, le FCR doit être calculé, bien qu'il ne s'agisse pas d'un élément

critique pour l'examen officiel des mesures de la qualité statistique dans le cadre de l'enquête GATS.

2. **Estimation du FCR** □□ Le numérateur du FCR peut être estimé en calculant la somme des pondérations (de base) avant ajustement sur tous les ménages sélectionnés. Le dénominateur doit être obtenu auprès d'une source à jour et hautement réputée (par exemple, un recensement récent ou une autre enquête récente dont l'échantillon est d'une taille supérieure à celle de l'enquête GATS).
3. **Rapport du FCR** □□ Le FCR estimé doit figurer dans le *Rapport national de l'enquête GATS* accompagné de la description du plan d'échantillonnage ou d'une évaluation indépendante de la qualité des données.
4. **Interprétation du FCR** □□ Le FCR indique l'ampleur de la couverture de la base de sondage des ménages, mais pas la taille du biais de couverture, puisque cette dernière dépend de la première, à laquelle on ajoute la différence cumulée (dans les principales mesures de l'enquête) entre les ménages couverts et non couverts. Un taux de couverture des ménages supérieur à 95 % est considéré comme acceptable.
5. **Exemple** □□ La meilleure estimation extérieure disponible du nombre de ménages au sein de la population cible d'un pays participant à l'enquête GATS est de 36 436 103, alors que la somme pondérée des 13 000 ménages sélectionnés pour l'échantillon de l'enquête GATS est de 34 788 543. Le taux de couverture résultant de $0.955 = \frac{34,788,543}{36,436,103}$ serait considéré comme acceptable.

Un taux de couverture des ménages supérieur à 95 % est considéré comme acceptable.

5.3.2 Tendances de taux d'abandon des répondants

Certains entretiens d'enquête ne sont pas menés à leur terme. Le répondant peut trouver l'entretien pesant et perdre patience avec l'enquêteur, ou peut avoir besoin de mettre fin à l'entretien plus tôt en raison de contraintes d'emploi du temps légitimes et ne pas être en mesure ou désireux de reprogrammer un créneau pour terminer l'entretien.

Dans les deux cas, un entretien incomplet est un signe de désengagement des répondants. Que ce désengagement soit dû à l'approche adoptée par l'enquêteur ou bien à l'attitude du répondant à l'égard du sujet de l'enquête, il influe finalement sur la qualité des données.

- **Source(s) de données** : Le fichier de données contenant les codes de décision finale des ménages /individus doit être utilisé pour ces calculs (voir l'*annexe C.1*).
Un code de décision de 11 au *questionnaire individuel* indique que le répondant a rempli au moins les sections A à D de l'entretien individuel du GATS. Un code de décision de 12 au *questionnaire individuel* indique que l'entretien a été interrompu avant que les sections A à D puissent être réalisées.
- **Méthode/Calcul** : Définissez un « répondant » à l'enquête comme toute personne sélectionnée ayant reçu un code de décision finale de 11 ou 12 au *questionnaire individuel*. Définissez également le taux d'abandon (COR) comme :

$$COR = \frac{R_c}{R} = \frac{\# \text{ of GATS respondents who only partially completed the interview}}{\# \text{ of GATS respondents}}$$

ou

$$= \frac{\text{Code de décision 12}}{\text{Codes de décision (11 + 12)}}$$

- **Utilisations** : Les valeurs du taux d'abandon de l'échantillon doivent être calculées directement lors de la phase finale du processus d'assurance qualité, après la collecte des données. Par ailleurs, les superviseurs de terrain doivent s'efforcer d'identifier les enquêteurs de terrain inefficaces et prendre des mesures pour y remédier, tout au long de la collecte de données (en utilisant les codes de résultat de terrain correspondants pour le calcul). Cela implique de calculer les taux d'abandon par enquêteur individuel ou par équipe d'enquêteurs, et d'identifier les enquêteurs/équipes d'enquêteurs qui affichent des taux particulièrement élevés (par exemple des taux supérieurs à la moyenne + deux écarts types des taux parmi tous les enquêteurs/équipes). Les valeurs de taux d'abandon peuvent également être calculées selon : 1) la semaine de collecte de données pendant laquelle l'entretien a eu lieu, 2) l'âge du répondant, 3) le genre du répondant et 4) le milieu urbain/rural du répondant.
- **Interprétation** : En règle générale, il est préférable que la valeur du taux d'abandon soit la plus faible possible. Une baisse du taux d'abandon sur l'ensemble de la durée de collecte des données pourrait indiquer que les répondants étaient moins enclins à fournir des données utiles à l'enquête GATS, vers la fin de collecte de données. Bien que les taux d'abandon constituent d'intéressantes mesures du rendement de la collecte de données et de la qualité globale de l'enquête, il n'est pas essentiel de les prendre en compte dans le cadre de l'examen officiel des mesures de la qualité statistique de l'enquête GATS.

5.3.3 Tendances des taux de réponse des ménages par strate d'échantillonnage de la première phase

Le taux de réponse des ménages est calculé comme le rapport entre le nombre pondéré de ménages pour lesquels une liste complète de données a été fournie et le nombre pondéré de ménages dans la zone (voir la formule 19 fournie dans le *Manuel de pondération de l'échantillon de l'enquête GATS*). Les ménages qui ne fournissent qu'une liste de données partielle sont traités comme des refus ou des interruptions pour le taux de réponse des ménages, et ne sont pas inclus dans le numérateur du calcul du taux de réponse des ménages. Le *questionnaire destiné aux ménages* est considéré comme complet s'il a un code de décision finale de 1 (voir l'*annexe C.1*). En considérant uniquement les ménages de la zone de l'enquête, le taux de réponse des ménages est calculé comme suit :

$$\text{Household - Level RR} = \frac{[1]}{[1] + [3] + [4] + [5] + [6] + [9]} \times 100$$

où

1 = Questionnaire destiné aux ménages rempli ; une personne sélectionnée

2 = Questionnaire destiné aux ménages rempli ; aucune personne sélectionnée

- 3 = Questionnaire destiné aux ménages partiellement rempli ; liste interrompue (entretien incomplet)
- 4 = Questionnaire destiné aux ménages non rempli ; aucun répondant au questionnaire de sélection des ménages n'a pu être identifié
- 5 = Aucune personne présente au domicile
- 6 = Le ménage n'a pas souhaité répondre au questionnaire
- 9 = Autre motif de non-réponse du ménage.

Notez que le code de décision finale 2 au *questionnaire destiné aux ménages* est exclu à la fois du numérateur et du dénominateur du taux de réponse des ménages, puisque ces ménages ne sont pas considérés comme admissibles. Ce taux de réponse est conforme à RR1, tel que défini par l'Association américaine de recherche sur l'opinion publique (AAPOR) (<https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx>).

Une documentation complète décrivant le calcul des taux de réponse des ménages doit être fournie par les pays. Les taux de réponse des ménages doivent être présentés pour l'ensemble du pays selon les strates d'échantillonnage de la première phase (c'est-à-dire par milieu urbain/rural et par région) si la taille des échantillons est suffisante pour une estimation régionale. Une ventilation supplémentaire, si nécessaire, doit être effectuée en utilisant la classification croisée par milieu urbain/rural et par région.

5.3.4 Tendances des taux de réponse au niveau individuel parmi les variables utilisées pour les ajustements pour non-réponse

Un taux de réponse individuel correspond au rapport entre le nombre pondéré d'entretiens GATS partiellement ou totalement achevés et le nombre pondéré d'individus admissibles échantillonnés choisis dans la liste de tous les membres du ménage. Concrètement, les codes de décision finale pour les individus sont les suivants :

- 11 = Questionnaire individuel rempli
- 12 = Entretien incomplet
- 13 = La personne sélectionnée s'est ultérieurement avérée non admissible à l'enquête GATS
- 14 = Le répondant sélectionné était absent du domicile
- 15 = Le répondant sélectionné n'a pas souhaité répondre
- 16 = Le répondant sélectionné n'était pas qualifié
- 17 = Autre motif de non-réponse.

L'une des manières de calculer le taux de réponse individuelle est la suivante :

$$Person - Level RR = \frac{[11]}{[11] + [12] + [14] + [15] + [16] + [17]} \times 100$$

Le taux de réponse individuelle ci-dessus présume que tout individu sélectionné dont l'admissibilité est inconnue (par exemple avec un code de décision finale de 14) est admissible pour répondre à l'enquête GATS. Cela peut conduire à sous-estimer les taux de réponse individuelle si les enquêteurs choisissent régulièrement des répondants qui, une fois l'entretien entamé, s'avèrent inadmissibles à l'enquête. Il est

recommandé aux pays d'estimer la proportion de répondants sélectionnés à partir de la liste qui sont réellement admissibles pour répondre à l'enquête GATS (e) en utilisant des fréquences pondérées de codes de décision telles que :

$$e_{\text{person-level}} = \frac{[11] + [12] + [15] + [16] + [17]}{[11] + [12] + [13] + [15] + [16] + [17]}$$

Si cette proportion pour l'ensemble de l'échantillon est inférieure à 0,90, les pays doivent ajuster la composante inconnue du taux de réponse individuelle en multipliant les inconnues (code de décision finale de 14) par ce rapport (e). La formule suivante est conforme au taux de réponse RR3 de l'AAPOR (2004) :

$$\text{Person - Level RR} = \frac{[11]}{[11] + [12] + [e_{\text{person-level}}[14]] + [15] + [16] + [17]} \times 100$$

Ce taux de réponse est conforme à RR3 tel que défini par l'AAPOR et documenté ci-dessus.

Une documentation complète décrivant le calcul des taux de réponse doit être fournie par les pays.

Les taux de réponse individuelle doivent être présentés pour l'ensemble du pays et par sous-zones au sein de la région lorsque la taille des échantillons est suffisante pour une estimation régionale. Les variables utilisées pour définir l'ajustement pour non-réponse doivent donner lieu à une ventilation supplémentaire au sein de la région, le cas échéant. Cette ventilation peut tenir compte du milieu de vie (urbain/rural), de l'âge indiqué dans la liste (15-24 ans, 25-34 ans, 35-44 ans, 45-54 ans, 55 ans ou plus), du genre (homme/femme) et du statut tabagique actuel (fumeur/non-fumeur).

Le *Manuel GATS de supervision sur le terrain* décrit les responsabilités des superviseurs de terrain concernant l'examen et la prise de mesures correctrices dans les cas où une équipe d'enquêteurs affiche de faibles taux de réponse.

5.3.5 Tendances des taux de refus au niveau individuel parmi les variables utilisées pour les ajustements pour non-réponse

Les taux de refus sont calculés comme le rapport entre le nombre pondéré de répondants admissibles sélectionnés qui ont refusé de répondre au *questionnaire individuel* de l'enquête GATS (code de décision finale de 15 au *questionnaire individuel*) et le nombre total pondéré de questionnaires individuels remplis (11) et incomplets (12) et de non-répondants dont l'admissibilité n'est pas connue (14), d'individus n'ayant pas souhaité répondre (15), de répondants non qualifiés (16) et d'autres individus non-répondants (17).

En utilisant les codes de décision au niveau individuel, le calcul est le suivant :

$$\text{Person - Level Refusal Rate} = \frac{[15]}{[11] + [12] + [14] + [15] + [16] + [17]} \times 100$$

Ce taux de refus est conforme à REF1 tel que défini par l'Association américaine de recherche sur l'opinion publique (AAPOR).

Les taux de refus individuels de l'échantillon doivent être calculés directement lors de la phase finale du processus d'assurance qualité, après la collecte de données. Par ailleurs, les superviseurs de terrain doivent s'efforcer d'identifier les enquêteurs de terrain inefficaces et prendre des mesures pour y remédier, tout au long de la collecte de données. Cela implique de calculer les taux de refus par enquêteur individuel ou par équipe d'enquêteurs, et d'identifier les enquêteurs/équipes d'enquêteurs qui affichent des taux particulièrement élevés (par exemple des taux supérieurs à la moyenne + deux écarts types des taux parmi tous les enquêteurs/équipes). Ces taux de refus peuvent également être calculés selon : 1) la semaine de collecte de données pendant laquelle l'entretien a eu lieu, 2) l'âge du répondant, 3) le genre du répondant et 4) le milieu urbain/rural du répondant.

Une documentation complète décrivant le calcul des taux de refus doit être fournie par les pays. Les taux de refus doivent être présentés pour l'ensemble du pays et par sous-zones au sein de la région si la taille des échantillons est suffisante pour une estimation régionale. Les variables utilisées pour définir l'ajustement pour non-réponse doivent donner lieu à une ventilation supplémentaire au sein de la région, le cas échéant. Cette ventilation peut tenir compte du milieu de vie (urbain/rural), de l'âge indiqué dans la liste (15-24 ans, 25-34 ans, 35-44 ans, 45-54 ans, 55 ans ou plus), du genre (homme/femme) et du statut tabagique actuel (fumeur/non-fumeur).

5.3.6 Taux de non-réponse par question pour les variables d'indicateurs de la fiche d'information

Un taux de non-réponse à une question de l'enquête GATS (par exemple, le pourcentage de tous les répondants qui ne parviennent pas à fournir de données utiles à une question spécifique du questionnaire de l'entretien sur l'ensemble des répondants censés répondre à la question) doit être calculé et documenté après le traitement des données de l'enquête pour les éléments du questionnaire qui sont associés à tous les indicateurs qui seront inclus dans la *Fiche d'information de l'enquête GATS* spécifique à chaque pays (voir les *Définitions des indicateurs de l'enquête GATS* pour plus d'informations).

Les taux de non-réponse à une question correspondent au rapport entre le nombre de répondants qui n'ont pas fourni de réponse valide à une question prévue (*RÉPONSE MANQUANTE^x* à la question x) et le nombre total de répondants au niveau des unités (I) moins le nombre de répondants ayant passé la question x pour une raison valide (*V^x*) :

$$NRRI^x = \frac{MISSING^x}{(I - V^x)} \times 100$$

Le nombre total des non-répondants au niveau des unités pour x (*RÉPONSE MANQUANTE^x*) sera obtenu à partir d'une fréquence non pondérée des répondants avec des données manquantes pour l'élément x, après un nettoyage approprié destiné à s'assurer que les instructions « passer à » ont été correctement appliquées. Le nombre total de répondants au niveau des unités sera obtenu à partir de la fréquence totale non pondérée des répondants de sexe masculin ou féminin à la question A01 (une variable n'ayant pas de champs vides prévus). Le nombre total de répondants avec un enchaînement valide pour l'élément x peut être obtenu comme étant la fréquence de l'élément x avec une réponse en .S (voir la *section 5.1.3*). De nombreux indicateurs de la *Fiche d'informations* sont des mesures composites obtenues à partir de réponses à plusieurs questions. La non-réponse à toute question utilisée pour obtenir une mesure composite entraînera une non-réponse à la question pour la mesure composite. Les taux de non-réponse inférieurs à 5 % sont considérés comme faibles. Comme pour les mesures de

qualité présentées aux sections 5.3.2 et 5.3.5 ci-dessus, les superviseurs de terrain doivent s'efforcer de repérer continuellement les questions des enquêteurs ou des équipes d'enquêteurs obtenant un taux excessif de non-réponses, afin d'améliorer la qualité de la collecte des données au moment où celle-ci se déroule sur le terrain.

5.4 Examen formel de la qualité statistique

Un protocole scientifique a été élaboré afin que le Comité d'examen des échantillons de l'enquête puisse effectuer un examen officiel des mesures de la qualité statistique dans chaque pays participant à l'enquête GATS, après le traitement des données de l'enquête et le calcul des pondérations de l'échantillon, mais avant l'analyse des données et la production des résultats. Cet examen a pour objectif de vérifier que la pondération de l'échantillon a été calculée conformément aux spécifications du *Manuel de pondération de l'échantillon de l'enquête GATS*, d'examiner diverses mesures indiquant la qualité des données de l'enquête, et de s'assurer que les principaux aspects de l'échantillonnage et de l'analyse sont documentés de manière exhaustive. Une assistance technique sera fournie pour la préparation de cet examen.

Les mesures de qualité spécifiques ainsi que la documentation nécessaires pour effectuer ces examens sont décrites ci-après.

5.4.1 Documentation de la pondération de l'échantillon

Chaque étape de la pondération des données et du processus d'ajustement doit être intégralement décrite dans un document préparé par les statisticiens du pays directement responsables du calcul et de l'ajustement de la pondération statistique de l'échantillon GATS. Ce document doit comporter un texte décrivant en détail le processus de pondération et d'ajustement, avec des exemples illustrant précisément la manière dont toutes les composantes de la pondération finale ont été calculées pour 5 à 10 répondants de l'échantillon réel. Plus spécifiquement, chacun des éléments suivants doit être préparé pour être étudié par les personnes chargées de l'examen¹ :

- Résumez le plan d'échantillonnage de l'enquête par étape et décrivez la manière dont les probabilités de sélection de chaque étape et la pondération de base globale ont été calculées.
- Décrivez en détail la manière dont les stratégies d'ajustement pour non-réponse et les stratégies d'ajustement d'étalonnage ont été exécutées, en précisant spécifiquement : i) comment les catégories de pondération ont été définies concernant les ajustements pour non-réponse à l'échelle des ménages et des individus, ii) quelles cellules d'ajustement ont été créées pour étalonner la pondération, et iii) sur quels effectifs de population la pondération a été étalonnée. Reportez-vous au *Manuel de pondération de l'échantillon*, qui vous orientera dans la rédaction de cette description.
- Choisissez 5 à 10 répondants réels pour illustrer le processus de calcul des pondérations qui a été suivi pour ces personnes. Ces répondants doivent être sélectionnés dans différentes strates, UPE et catégories de genre. Si des éléments de pondération fondés sur des intervalles semi-ouverts ont été utilisés, un répondant de l'un de ces groupes géographiques doit alors être inclus

¹ Puisque de nombreux éléments de cette liste doivent également être inclus dans la documentation de chaque *Rapport national de l'enquête GATS*, notez que les ébauches de texte de cette documentation (mais contenant les versions finales des résultats numériques requis) peuvent être envoyées aux personnes qui effectuent l'examen officiel de la qualité statistique pour le pays.

dans l'un des 5 à 10 exemples. À l'issue de cette étape, le calcul de la pondération finale ajustée doit être affiché pour chaque exemple. Plus précisément, pour chacun de ces répondants :

- Identifiez les composantes de chaque probabilité conditionnelle et inconditionnelle (à partir de chaque étape de l'échantillonnage) qui ont été utilisées pour produire la PONDÉRATION DE BASE.
 - Présentez et décrivez intégralement chaque composante du calcul de l'AJUSTEMENT POUR NON-RÉPONSE à l'échelle des ménages et des individus. À l'issue de cette étape, la pondération de base doit être ajustée pour la non-réponse des ménages et des individus.
 - Présentez et décrivez de manière exhaustive chaque étape de la méthode D'AJUSTEMENT DE LA PONDÉRATION D'ÉTALONNAGE utilisée pour s'assurer que la distribution finale pondérée de l'échantillon correspond à la population d'étalonnage choisie.
- Présentez les résultats de calcul suivants basés sur les formules fournies aux *sections 5.2 et 5.3* de ce manuel :
 - Ajustements d'étalonnage par stratification a posteriori pour chaque cellule d'ajustement utilisée pour l'étalonnage (5.2.1)
 - Effet multiplicateur de la variable de pondération de l'échantillon (*Meff*) par tous les sous-groupes de population pour lesquels des estimations de l'enquête seront communiquées dans le *Rapport national de l'enquête GATS* (5.2.2)
 - Mesures individuelles et synthétiques (à savoir, minimum, maximum et médiane) de l'effet du plan de sondage global et la mesure de l'homogénéité au sein d'une UPE pour toutes les estimations de l'enquête présentées dans la *Fiche d'information de l'enquête GATS* (5.2.3)
 - Les mesures individuelles et synthétiques (à savoir, minimum, maximum et médiane) de la marge d'erreur, ou d'une autre mesure comparable de la précision, pour toutes les estimations de l'enquête présentées dans la *Fiche d'information de l'enquête GATS* (5.2.4)
 - Mesures individuelles et synthétiques (à savoir, minimum, maximum et médiane) des taux de réponse des ménages suivant les cellules d'ajustement utilisées en vue de produire l'ajustement pour la non-réponse au niveau des ménages (5.3.3)
 - Taux de réponse individuel parmi les cellules d'ajustement utilisées pour les ajustements pour la non-réponse au niveau individuel (5.3.4)
 - Taux de refus individuel parmi les cellules d'ajustement utilisées pour les ajustements pour la non-réponse au niveau individuel (5.3.5)
 - Mesures individuelles et synthétiques (à savoir, minimum, maximum et médiane) des taux de non-réponse pour tous les éléments de l'enquête utilisés en vue de calculer les taux de consommation de tabac présentés dans la *Fiche d'information de l'enquête GATS* (5.3.6)

Le *Chapitre 4* du *Manuel de pondération de l'échantillon de l'enquête GATS* suggère plusieurs méthodes pour assurer au mieux la qualité de la pondération de l'échantillon. Les personnes chargées de l'examen de la qualité statistique réaliseront les vérifications suivantes (les données nécessaires à la réalisation de ces vérifications devront donc être obtenues) :

- La taille moyenne de la pondération de base, divisée par la taille moyenne de la pondération ajustée pour la non-réponse, doit être approximativement égale au taux de réponse individuelle global final.
- La répartition pondérée de l'échantillon (en utilisant la pondération finale ajustée) dans toutes les cellules d'ajustement d'étalonnage doit refléter précisément la répartition correspondante des effectifs de la population extérieure utilisés pour l'étalonnage des pondérations.
- La somme des pondérations finales ajustées sur l'ensemble de l'échantillon doit correspondre à l'effectif total de la population (âgée de 15 ans ou plus) de la source utilisée pour l'étalonnage. Si cette source correspond au dernier recensement en date, alors la somme des pondérations doit correspondre aux effectifs de la population cible selon le dernier recensement.
- La plupart des ajustements par stratification a posteriori (par cellule d'ajustement) doivent être légèrement supérieurs ou inférieurs à 1.
- Le *Meff* doit être $\leq 2,00$ pour chaque sous-groupe rapporté dans le *Rapport national de l'enquête GATS*.

5.5 Création du fichier de données analytiques

Une fois la pondération de l'échantillon et tous les contrôles de qualité réalisés, un nouveau fichier doit être créé, contenant seulement les enquêtes avec un code de décision finale individuelle de 11 (Questionnaire *individuel rempli*). Seules les enquêtes ayant un code de décision finale individuelle de 11 seront classées comme des « répondants » au *Questionnaire individuel* de l'enquête GATS.

Notez qu'il n'est pas nécessaire de créer l'ensemble de données analytiques final avant de procéder aux contrôles d'assurance qualité en utilisant la marge d'erreur, l'effet du plan de sondage ou les principales estimations de l'enquête calculés par les procédures SAS ou SPSS pour des échantillons complexes, étant donné que les enquêtes avec des codes de décision finale autres que 11 au questionnaire individuel n'auront pas de valeurs pour la variable de pondération (à savoir, que ces valeurs seront manquantes) et seront automatiquement omises par le logiciel en tant que défauts, sauf indication contraire.

Ce nouveau fichier, que l'on appelle fichier de données analytiques, doit être utilisé lors de l'analyse des données afin de produire les estimations de l'enquête (c'est-à-dire des indicateurs), y compris les estimations qui figurent dans la *Fiche d'information* et le *Rapport national* de l'enquête GATS.

5.6 Création du fichier de données destiné à une utilisation publique

L'ensemble de données destinées au grand public est le fichier de données rendu public par le CDC. Ce fichier comporte uniquement les entretiens complets, avec toutes les variables de réponse de l'enquête

ainsi que les principales variables d'indicateurs normalisées de l'enquête GATS. Ce fichier de données contient également les variables géographiques ainsi que les variables de pondération de l'échantillon telles que la strate, le groupe et les pondérations finales de l'enquête.

L'ensemble de données destinées au grand public est publié pour que les données de l'enquête GATS soient accessibles à d'autres chercheurs. Conformément à la *Politique de publication des données de l'enquête GATS*, ce fichier est publié un an après que le pays a rendu publique sa *Fiche d'information de l'enquête GATS*.

6. Critères en matière de données et de documentation

Ce chapitre décrit les fichiers de données et la documentation qui seront envoyés par les pays et archivés au Centre de coordination des données de l'enquête GATS.

Fichiers de données et documents demandés au pays :

1. Fichiers DB3 finaux programmés : CMSDB.db3, Survey0.db3 et Survey1.db3.
2. Fichier maître DB3 créé par le logiciel d'agrégation à partir de CMSDB.db3, Survey0.db3 et Survey1.db3.
3. Fichier de base de données maître (avec variable de langue anglaise), généré à l'aide de la fonction « Création des fichiers d'analyse » contenant les fichiers plats transposés pour la table de réponses (avec des réponses valides au *questionnaire destiné aux ménages* et au *questionnaire individuel*) ainsi que le fichier de sélection de l'échantillon-maître (numéros d'enquête et identifiants de l'échantillon).
4. Fichiers de syntaxe SAS, SPSS et STATA.
5. Composantes et documentation sur la pondération de l'échantillon (se référer à la *section 5.4*).
6. Ensemble de données pondéré : les fichiers finaux, plats, nettoyés et pondérés, ainsi que les spécifications de pondération.

7. Bibliographie

- Biemer, Paul P. et Lars E. Lyberg, *Introduction to Survey Quality*. John Wiley and Sons, Inc., Hoboken, New Jersey, 2003.
- Encyclopedia of Survey Research Methods*, Lavrakas, P.J. (Ed.), Sage Publishing, Inc. Los Angeles, États-Unis d'Amérique d'Amérique, 2007.
- Gabler, S., Haeder, S., et Lahiri, P., « A model based justification of Kish's formula for design effects for weighting and clustering », *Survey Methodology*, 25 (1), 1999, p. 105-106.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Spécifications pour la programmation du questionnaire principal*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Questionnaire principal avec questions facultatives*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Plan de mise en œuvre pour la gestion des données*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Guide de formation à la gestion des données*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Manuel d'enquête sur le terrain*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Manuel de supervision sur le terrain*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Manuel de cartographie et de constitution des listes*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Guide de programmation sur General Survey System*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Spécifications question par question*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Manuel d'échantillonnage*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Groupe de collaboration sur l'enquête mondiale sur le tabagisme chez les adultes, *Enquête mondiale sur le tabagisme chez les adultes (GATS) – Manuel de pondération de l'échantillon*. Atlanta, Géorgie (États-Unis d'Amérique) : Centers for Disease Control and Prevention, 2019.
- Référence interne pour les valeurs manquantes dans SAS :
<http://support.sas.com/documentation/cdl/en/Ircon/61722/HTML/default/a000992455.htm>
- Kish, L., *Survey Sampling*, Wiley and Sons, New York, 1965.
- Kalton, G. et Flores-Cervantes, I., « Weighting Methods », *Journal of Official Statistics*, 19(2), 2003, p. 81-97.

Potter, F. J., *Survey of Procedures to Control Extreme Sampling Weights. Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1988, p. 453-458.

Annexe A : Glossaire

Dossier d'enquête : La liste et les informations de l'échantillon qui sont chargées sur les dispositifs portables pour la collecte de données.

Questionnaire principal : L'ensemble des questions de base posées dans le cadre de l'enquête GATS. Les pays participants adaptent le questionnaire principal à leurs besoins spécifiques en suivant des directives précises.

Fichiers DB3 : La programmation du logiciel d'enquête GSS se fait en utilisant des fichiers DB3 (SQLite).

General Survey System (GSS) : Le logiciel d'enquête qui est chargé sur les dispositifs individuels afin de réaliser l'enquête GATS.

Spécifications de programmation sur GSS : Outre le questionnaire papier qu'un pays utilise pour créer sa propre version adaptée, des spécifications de programmation sur GSS sont élaborées pour chaque version adaptée au contexte national et sont utilisées afin de programmer le questionnaire sur les dispositifs portables.

Configuration des dispositifs portables : Le processus de copie du logiciel sur le dispositif portable, généralement effectué sur tous les dispositifs portables en même temps avant le début de la formation des enquêteurs de terrain.

Fichier de sélection de l'échantillon-maître : L'ensemble de données qui contient un numéro d'enquête pour chaque ménage figurant dans la liste d'échantillonnage ainsi que les informations nécessaires pour calculer la pondération de l'échantillon et analyser les données complexes de l'enquête.

Assurance qualité : Un processus consistant en diverses activités systématiques visant à assurer, évaluer et confirmer la qualité des données collectées lors d'une enquête.

Comité d'examen du questionnaire (Questionnaire Review Committee, QRC) : Un groupe composé d'experts internationaux de la lutte antitabac et de la conception de questionnaires. Le QRC examine et approuve tous les questionnaires de l'enquête GATS adaptés aux pays afin d'en assurer la qualité, la normalisation et la comparabilité.

Comité d'examen des échantillons de l'enquête (Sample Review Committee, SRC) : Un groupe composé d'experts internationaux en matière d'échantillonnage. Le SRC examine et approuve les plans d'échantillonnage, la pondération des échantillons, ainsi que les mesures d'assurance qualité des données des pays participants.

Annexe B : Actions à mener avant la collecte des données

B.1 Adaptation du questionnaire GATS et processus d'examen

Adaptation du questionnaire : En collaboration avec le CDC, les pays participant à l'enquête GATS travailleront à l'élaboration d'un questionnaire GATS adapté à leur contexte national. Cette adaptation couvre les aspects suivants :

- **Adaptation du questionnaire principal de l'enquête GATS.** Chaque pays adaptera le questionnaire principal à sa situation spécifique, en modifiant les listes de questions (par exemple au sujet des types de tabac) et les réponses possibles (par exemple la liste des marques de cigarettes) selon le contexte national.
- **Ajout de questions facultatives.** L'enquête GATS comporte une liste de questions facultatives que les pays peuvent choisir d'ajouter ou non dans leur questionnaire.
- **Ajout de questions supplémentaires élaborées par le pays.** Les pays ont la possibilité d'ajouter leurs propres questions pour aborder des aspects qui ne figurent pas dans le questionnaire principal ou dans la liste de questions facultatives.
- **Suppression des questions non pertinentes.** Les pays ont la possibilité de retirer des questions principales qui ne sont pas adaptées à leur contexte (par exemple si la prévalence du tabac sans fumée est nulle). Une justification appropriée doit être fournie.

L'examen du questionnaire GATS est un processus en dix étapes :

1. **Adaptation du projet de questionnaire au contexte national.** En collaboration avec le CDC, les pays travailleront à l'élaboration d'un projet de questionnaire adapté au contexte national, à soumettre au Comité d'examen du questionnaire.
2. **Mise en forme.** Le CDC examinera le projet de questionnaire soumis et le mettra en forme au besoin (en modifiant par exemple des instructions « passer à » ou en mettant en relief certains aspects en vue de l'examen du Comité d'examen du questionnaire).
3. **Examen d'experts par le Comité d'examen du questionnaire.** Les membres du Comité d'examen du questionnaire examinent le questionnaire et formulent des commentaires.
4. **Compilation des commentaires des examinateurs.** Le CDC compile les commentaires du Comité d'examen du questionnaire et les renvoie au pays pour examen.
5. **Communication (résolution).** Le QRC et le pays travaillent ensemble à la résolution des questions en suspens (par courriel et/ou conférences téléphoniques).
6. **Finalisation du questionnaire adapté.** Une fois que tous les problèmes sont résolus, le questionnaire adapté au pays est finalisé et l'approbation officielle est donnée par le QRC.
7. **Notification d'acceptation par le pays.** Le pays accorde son acceptation finale du questionnaire approuvé par le QRC.
8. **Traduction.** Les pays font traduire la version finale approuvée du questionnaire GATS dans la ou les langues nécessaires et fournissent une rétro-traduction du questionnaire en anglais.
(Remarque : La traduction peut être réalisée en même temps que l'étape 1 : dans ce cas, les

pays traduisent d'abord le questionnaire principal et l'adaptent dans leur propre langue, puis le retraduisent en anglais pour le soumettre au QRC).

9. **Vérification des versions finales traduites et rétro-traduites.** Le QRC examinera les versions traduites et rétro-traduites afin de vérifier l'exactitude de la traduction (en veillant à ce que le sens des questions ait bien été conservé). (Remarque : Cette démarche fait partie du processus d'examen du QRC si la traduction est réalisée en même temps que l'étape 1 et qu'une version de traduction inversée est initialement soumise pour examen au QRC).
10. **Programmation.** Des spécifications de programmation sont élaborées pour le questionnaire adapté au pays ; l'enquête adaptée au contexte national est programmée sur les dispositifs portables.

B.2 Processus de programmation du questionnaire de l'enquête GATS

Les principales étapes et les recommandations relatives aux échéances de la programmation des questionnaires de l'enquête GATS sont les suivantes :

1. La proposition de questionnaire et de test préliminaire de l'enquête GATS est approuvée 6 à 8 semaines avant la formation au test préliminaire.
2. La programmation du questionnaire en anglais commence 6 semaines avant la formation au test préliminaire.
3. Les équipements portables sont livrés et opérationnels 4 semaines avant la formation au test préliminaire.
4. La traduction des menus et des messages du logiciel GSS, le cas échéant, est terminée 1 semaine avant la formation au test préliminaire.
5. La traduction des textes du questionnaire (questionnaire destiné aux ménages et questionnaire individuel) dans toutes les langues du pays d'accueil a été réalisée et intégrée dans les bases de données survey0 et survey1 deux semaines avant la formation au test préliminaire.
6. Le *questionnaire destiné aux ménages* et le *questionnaire individuel* sont approuvés par le pays d'accueil 1 semaine avant la formation au test préliminaire.
7. À compter de cette étape, les nouvelles versions du questionnaire doivent être strictement contrôlées par rapport à la version approuvée.

B.3 Processus de configuration du programme GSS sur les dispositifs portables

Cette section décrit les mesures de qualité qui doivent être mises en œuvre aux différents stades du processus afin de garantir que chaque dispositif portable envoyé sur le terrain possède bien le même logiciel et les mêmes spécifications matérielles.

1. **Création de la version maîtresse :** La version maîtresse doit être finalisée sur un seul ordinateur, exempt de tout virus et équipé d'un logiciel antivirus adapté. Une fois créée, la version maîtresse doit être chargée sur deux dispositifs portables différents et un essai complet doit être effectué sur les deux dispositifs, en répondant à chaque question du questionnaire. Si des erreurs sont détectées sur l'un des dispositifs portables, la programmation doit être corrigée

(comme il est mentionné dans la *section 3.3.2*), et ce processus doit être répété jusqu'à ce que l'exactitude soit de 100 %. Ce processus doit être réalisé en présence du coordinateur/de la personne référente et du responsable informatique.

2. **Application d'autocollants des vérifications de contrôle qualité (modèle) :** Un autocollant indiquant l'identifiant de l'enquêteur de terrain, le numéro de série, ainsi qu'une liste des vérifications de contrôle qualité à effectuer doit être placé à l'arrière de chaque dispositif portable. Voir l'exemple ci-dessous :

701010	SN A10	
	QC1	QC2
☐ Config initiale	☐ Date	☐ Date
☐ APK Loaded	☐ FIID (identifiant d'enquêteur de terrain)	☐ FIID (identifiant d'enquêteur de terrain)
☐ Transmission	HQ _____ (Questionnaire auprès des ménages)	☐ Delete Training Cases (Effacer les enquêtes de formation)
	QI _____ (Questionnaire individuel)	☐ Alimentation _____
Ini _____	☐ Alimentation _____	Transmission
	Transmission	Enquêtes _____
	QC1 _____	QC2 _____

3. **Préparation des dispositifs portables :** Une fois la version finale copiée, tous les dispositifs portables doivent être configurés étape par étape en suivant la procédure appropriée (voir la procédure de préparation des dispositifs portables décrite dans le *Guide de programmation sur General Survey System*). Le processus de répartition des dispositifs portables dans différentes boîtes (avant et après configuration) doit être appliqué à nouveau.
4. **Réalisation de la première phase de contrôle qualité (QC1) :** Les dispositifs portables doivent être équitablement répartis entre les différentes personnes chargées du contrôle qualité, qui doivent faire partie de l'équipe ayant configuré les dispositifs portables ou supervisé le processus. Chaque étape de la première phase de contrôle qualité doit être réalisée sur tous les dispositifs de poche. La première phase de contrôle qualité doit être un processus rigoureux de vérification de la qualité. L'exactitude des éléments suivants doit être contrôlée sur chaque dispositif portable, puis cochée sur l'autocollant où figure la liste de contrôle :

- 1) Date et heure
- 2) Identifiant de l'enquêteur de terrain (vérifier l'autocollant avec l'identifiant du *questionnaire destiné aux ménages*)
- 3) Version du questionnaire destiné aux ménages
- 4) Version du questionnaire individuel
- 5) Alimentation – autonomie de la batterie

D'autres contrôles peuvent être ajoutés en fonction du contexte du pays, comme le nombre d'enquêtes et de dispositifs portables pour les enquêteurs hommes/femmes, etc. Pour tout

problème identifié, le dispositif portable concerné doit être placé dans un carton différent et être réinitialisé à nouveau, puis l'ensemble du processus doit être répété sur ce dispositif portable à partir de l'étape 1.

5. **Réalisation de la seconde phase de contrôle qualité (QC2) :** La seconde phase de contrôle qualité est effectuée une fois que tous les dispositifs portables ont été soumis à la première phase de contrôle qualité et que tous les points de la liste de contrôle figurant sur leur étiquette ont été validés. Les dispositifs portables doivent être redistribués de sorte que les première et seconde étapes de contrôle qualité soient réalisées par deux personnes différentes. La seconde phase de contrôle qualité reprend le même processus que la première phase ; l'étape supplémentaire propre à la seconde phase consiste à vérifier si tous les autocollants ont bien été cochés et validés. Il est important de réaliser la seconde phase de contrôle qualité, puisqu'il a été observé qu'un petit nombre d'erreurs ne sont détectées que lors de cette phase. Tout dispositif portable présentant des problèmes identifiés lors de la seconde phase de contrôle qualité doit être de nouveau soumis au processus de vérification, en commençant par une réinitialisation matérielle à l'étape 1.
6. **Validation de la fiche d'achèvement par le responsable informatique et le référent national :** Une fois que les première et seconde phases de contrôle qualité (QC1 et QC2) ont été réalisées sur chacun des dispositifs portables, un bref rapport de synthèse indiquant le nombre de dispositifs portables, les numéros de version du *questionnaire destiné aux ménages* et du *questionnaire individuel*, la date, l'heure ainsi que les membres de l'équipe chargée de la configuration et du contrôle qualité des dispositifs portables doit être documenté et validé par le responsable informatique et la personne référente de l'enquête GATS. Cette validation doit indiquer que les dispositifs portables sont prêts pour la formation ou le travail de terrain de l'enquête complète et qu'aucun autre changement de programmation ne sera effectué. Cette validation signifie également que les dispositifs portables ne seront utilisés à aucune autre fin que celle de l'enquête GATS, jusqu'à l'achèvement de l'enquête dans le pays.

Annexe C : Actions à mener après la collecte des données

C.1 Codes de décision finale et calcul des taux de réponse

Après un examen attentif des enquêtes nationales auprès des ménages déjà réalisées au sujet du tabagisme ou autre et utilisant une méthode de collecte de données par téléphone ou en face à face, une liste de codes de décision finale et les calculs des taux de réponse ont été finalisés pour l'enquête GATS. Parmi les études examinées dans ce cadre figurent les enquêtes de l'Institut brésilien de géographie et de statistique (IBGE), l'enquête nationale de démographie et de santé sexuelle et reproductive de 2004 (ENDSSR-2004), l'enquête sur la santé reproductive de Géorgie (RHS), les enquêtes démographiques et sanitaires (DHS), l'enquête nationale sur la santé familiale (NFHS), le système de surveillance des facteurs de risques comportementaux (BRFSS), le système de surveillance des comportements à risque de la jeunesse (YRBSS), et l'Association américaine pour la recherche sur l'opinion publique (AAPOR).

Les codes de décision finale et les taux de réponse rapportés sont présentés dans les tableaux suivants de cette section. Les *tableaux C-1 et C-2* indiquent les codes de décision finale qui doivent être attribués pour différents codes de résultats définitifs au *questionnaire destiné aux ménages* et au *questionnaire individuel*. (De plus amples détails au sujet des codes de résultat sont disponibles dans le *Manuel d'enquête sur le terrain de l'enquête GATS*.) Le *tableau C-3* fournit un exemple de modèle pour le calcul et la notification des taux de réponse et des codes de décision dans le *Rapport national de l'enquête GATS*.

Tableau C-1 : Codes de décision finale et de résultat définitif du questionnaire destiné aux ménages

Codes de décision finale	Codes de résultats définitifs	Nom et description
1	200	Questionnaire destiné aux ménages rempli, une personne sélectionnée – Le questionnaire destiné aux ménages est considéré comme complet si la liste est complète
2	201	Questionnaire destiné aux ménages rempli, aucune personne sélectionnée – Personne n'est âgé de 15 ans ou plus – Personne ne considère le ménage comme son lieu de résidence habituel – Le ménage a été attribué à des hommes et aucune personne de sexe masculin ne réside dans le ménage ou le ménage a été attribué à des femmes et aucune personne de sexe féminin ne réside dans le ménage (si la randomisation par genre a été utilisée)
3	202	Questionnaire destiné aux ménages partiellement rempli ; liste non terminée – Entretien incomplet
4	203	Questionnaire destiné aux ménages incomplet ; aucune personne qualifiée identifiée pour répondre au questionnaire de sélection des ménages – Aucune personne âgée de 18 ans ou plus n'est présente – Membre du ménage âgé de 18 ans ou plus disponible mais non qualifié pour répondre – Il se peut qu'aucun résident de ces ménages ne soit admissible à l'enquête
5	209	Aucune personne présente au domicile
6	204	Le ménage n'a pas souhaité répondre au questionnaire
7	205	Le logement est inoccupé
8	206	L'adresse sélectionnée n'est pas le lieu de résidence d'un ménage
9	208	Autre motif de non-réponse du ménage

Tableau C-2 : Codes de décision finale et de résultat définitif du questionnaire individuel

Codes de décision finale	Codes de résultats définitifs	Nom et description
11	400, 402*	Questionnaire individuel rempli – Questionnaire rempli jusqu'à la question E01 – Aucune des questions essentielles (par exemple B01/B02/B03 et C01/C02/C03 le cas échéant) = Ne sait pas ou N'a pas souhaité répondre
12	402*	Incomplet – Questionnaire interrompu avant la question E01
13	403	La personne sélectionnée s'est ultérieurement avérée non admissible à l'enquête – Âge inférieur à 15 ans – La personne ne considère pas le ménage comme son lieu de résidence habituel – Erreur de genre (si la randomisation par genre a été utilisée)
14	409	Le répondant sélectionné était absent du domicile
15	404	Le répondant sélectionné n'a pas souhaité répondre
16	407	Le répondant sélectionné n'était pas en capacité de répondre
17	408	Autre motif de non-réponse

Les enquêtes dont le code de résultat définitif est 402 (questionnaire individuel partiellement rempli) se verront attribuer un code de décision de 11 (questionnaire individuel rempli) ou 12 (questionnaire individuel incomplet). Voir les règles définies à la *section 5.1.7*.

Tableau C-3 : Tableau 3.1 du rapport de pays de l'enquête GATS — Modèle de rapport des taux de réponse

Tableau 3.1 : Nombre et pourcentage de ménages et de personnes interrogées et taux de réponse, en fonction du lieu de résidence (non pondéré) – enquête GATS [pays], [année].						
	Lieu de résidence				Total	
	Urbain		Rural			
	Numéro	Pourcentage	Numéro	Pourcentage	Numéro	Pourcentage
<i>Ménage sélectionné</i>						
Entretien terminé (HC)						
Entretien terminé, mais aucune personne qualifiée pour répondre (HCNE)						
Entretien incomplet (HINC)						
Aucun répondant au questionnaire de sélection (HNS)						
Aucune personne présente au foyer (HNH)						
N'a pas souhaité répondre (HR)						
Le logement est inoccupé (HUO)						
L'adresse n'est pas un logement (HAND)						
Autre ¹ (HO)						
Nombre total de ménages sélectionnés		100		100		100
<i>Taux de réponse des ménages (HRR) (%)</i> ²						
<i>Personne sélectionnée</i>						
Entretien réalisé (PC)						
Entretien incomplet (PINC)						
Non éligible (PNE)						
Absent (PNH)						
N'a pas souhaité répondre (PR)						
N'était pas en capacité de répondre (PI)						
Autre ¹ (PO)						
Nombre total de personnes dans l'échantillon		100		100		100
<i>Taux de réponse individuel (PRR) (%)</i> ³						
<i>Taux de réponse total (TRR) (%)</i> ⁴						
¹ Tout autre résultat ne figurant pas dans la liste.			³ Le taux de réponse individuel (PRR) est calculé comme suit :			
² Le taux de réponse des ménages (HRR) est calculé comme suit :			PC *100 _____ PC + PINC + PNH + PR + PI + PO			
$\frac{HC \times 100}{HC + HINC + HNS + HNH + HR + HO}$			⁴ Le taux de réponse total (TRR) est calculé comme suit : (HRR x PRR) / 100			
Remarques :						
— Lorsqu'un entretien auprès d'un ménage est incomplet (c'est-à-dire que la liste de questions n'a pas pu être intégralement posée), ce dernier est considéré comme non-répondant à l'enquête GATS. Ces enquêtes (HINC) n'ont donc pas été prises en compte dans le numérateur du taux de réponse des ménages.						
— Le nombre total d'individus échantillonnés doit être égal au nombre d'entretiens auprès des ménages terminés [HC].						
— Un entretien individuel terminé [PC] inclut les répondants qui avaient répondu au moins à la question E01 et qui ont fourni des réponses valides aux questions B01/B02/B03 (et C01/C02/C03 le cas échéant). Les répondants qui ne remplissaient pas ces critères ont été considérés comme des non-répondants incomplets (PINC) à l'enquête GATS et n'ont donc pas été inclus dans le numérateur du taux de réponse individuel.						

C.2 Tendance des ajustements d'étalonnage des pondérations par stratification a posteriori au sein des cellules d'ajustement

Environnement de travail

La dernière étape de production des pondérations de l'échantillon consiste à calibrer les pondérations sur les effectifs de la population, en fonction de corrélats connus des principales mesures de résultats de l'étude, appelés variables d'étalonnage (par exemple, le sexe, le niveau d'études, l'âge, le milieu urbain/rural et la région, comme le suggère le *Manuel de pondération de l'échantillon de l'enquête GATS*). Selon la procédure, l'étalonnage

par stratification a posteriori implique la création de « cellules d'ajustement » par la classification croisée des mesures corrélées. L'ajustement par stratification a posteriori dans chacune de ces cellules d'ajustement est $< 1,00$ dans les catégories dont l'échantillon était surreprésenté après prise en compte de la sélection de l'échantillon et du taux de non-réponse, et $> 1,00$ dans les catégories où l'échantillon était sous-représenté. L'examen de la tendance d'ajustement par stratification a posteriori parmi les cellules d'ajustement permet de comprendre les déséquilibres restants au sein de l'échantillon après la correction du déséquilibre dû à la non-réponse.

Source(s) de données

Un fichier de données contenant l'ensemble des pondérations de l'échantillon final ajusté suivant le taux de non-réponse, ainsi que toutes les composantes de calcul de la pondération finale de chaque répondant, doit être utilisé pour ces calculs.

Logiciel de calcul

Aucun logiciel particulier n'est nécessaire, puisque les calculs à réaliser peuvent être aisément programmés dans n'importe quel logiciel.

Calcul

Les différentes étapes de la procédure de calcul des ajustements par stratification a posteriori sont décrites dans la *section 3.3 du Manuel GATS de pondération de l'échantillon*. En utilisant les résultats définitifs de ces calculs, organisez l'ensemble des ajustements par stratification a posteriori sous la forme d'un tableau multidirectionnel par rapport à la corrélation des variables qui définissent les cellules d'ajustement (à savoir, les variables d'étalonnage). L'idéal est que tous les ajustements par stratification a posteriori soient proches de 1,00, certains étant légèrement supérieurs à 1,00 et les autres légèrement inférieurs à 1,00.

Interprétation

Les écarts importants par rapport à 1,00 indiquent qu'il y a eu de graves déséquilibres de l'échantillon en raison de la couverture différentielle de l'échantillon des ménages, des problèmes non détectés de calcul des pondérations de l'échantillon, ou bien d'importantes évolutions démographiques de la population cible entre la date du recensement de la population à laquelle l'échantillon a été étalonné et la date de l'échantillonnage et de collecte des données de l'enquête GATS.

Tableau C-4 : Exemple fictif de tableau d'ajustements d'étalonnage des pondérations par stratification a posteriori, par région, milieu urbain/rural, genre et âge.

Région	Milieu urbain/rural	Genre	Âge	Ajustement
1	Urbain	Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
	Rural	Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
2	Urbain	Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
	Rural	Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
3	Urbain	Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
	Rural	Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Homme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	
		Femme	15-24 ans 25-44 ans 45-59 ans 60 ans ou plus	

C.3 Effet de la variable de pondération de l'échantillon sur la précision des estimations de l'enquête

Environnement de travail

La variation de pondération de l'échantillon crée souvent une augmentation de la variance (soit une réduction de la précision) des estimations de l'enquête. Ce type de variation constituera un problème significatif pour les plans d'échantillonnage de pays qui ont sur-échantillonné par zone urbaine/rurale ou par région. Kish (1965, *section 11.7*) a proposé un modèle simple, mais daté, pour mesurer l'augmentation multiplicatrice de la variance lorsque la taille de pondération n'est pas corrélée à la mesure de l'enquête utilisée pour estimer une caractéristique de population simple (θ). Nous utiliserons $Meff_w$ pour désigner cet effet multiplicateur.

$Meff_w$ n'est pas spécifique à l'estimation, mais il est spécifique au domaine de rapport des estimations. Il s'appliquera donc à toutes les estimations de l'enquête GATS dans le domaine concerné. Par exemple, $Meff_w$ calculé pour les estimations de la population globale doit être calculé à partir de l'échantillon entier, mais $Meff_w$ calculé pour les estimations rapportées des personnes vivant dans les zones rurales doit être calculé à partir de répondants des zones rurales seulement.

Source(s) de données

Le fichier final de données pondérées utilisé pour l'analyse doit être utilisé pour ces calculs.

Logiciel de calcul

Aucun logiciel spécifique n'est nécessaire. Le calcul de $Meff_w$ peut facilement être programmé en utilisant la sortie du logiciel d'analyse de l'enquête utilisé.

Calcul

En se concentrant sur la pondération finale (w_i) calculée pour chacun des n répondants auxquels $Meff_w$ s'applique, calculez la moyenne simple des pondérations pour obtenir $\bar{w}$, puis utilisez la

formule $s_w^2 = \sum_{i=1}^n (w_i - \bar{w})^2 / (n-1)$ pour calculer la variance des pondérations de l'échantillon. Nous pouvons calculer $Meff_w$ à partir de l'ensemble final des pondérations de l'échantillon des répondants de l'enquête GATS comme suit :

$$Meff_w = 1 + \frac{s_w^2}{\bar{w}^2}, \quad (1)$$

Notez que $Meff_w \geq 1$, puisque $\bar{w}^2$ et s_w^2 seront toujours positifs.

Interprétation

Alors qu'il est clairement préférable que $Meff_w$ soit aussi proche que possible de 1, $Meff_w > 2.00$ peut être considéré comme une valeur excessive appelant dès lors certaines mesures correctrices. La réduction de la pondération ou les stratégies de censure sont la solution la plus couramment utilisée pour

corriger les variables de pondération excessives (Potter, 1988). La décision finale de réduire la pondération dépend de l'obtention d'un équilibre entre l'avantage de la précision résultant de la réduction et de l'effet sur les estimations de l'enquête entraîné par les modifications de pondération. Si la réduction de la pondération fait baisser $Meff_w$, mais ne change pas sensiblement les estimations pondérées des résultats des principales mesures de l'étude, l'étape de réduction peut être justifiée, étant donné le temps nécessaire et les efforts à déployer.

Exemple :

Pour un échantillon de l'enquête GATS, la moyenne de pondération de l'échantillon final parmi les XXXX répondants à l'enquête est $\bar{w} = YYY.YYYY$, avec une variance simple de $s_w^2 = ZZZZ.ZZ$.

L'effet multiplicateur de la variable de pondération sur l'ensemble des estimations nationales d'un pays

$$\text{est } Meff_w = 1 + \frac{YYY.YYYY}{XXXX.ZZ} = U.UU$$

Un échantillon où $Meff_w = U.UU < 2.00$ peut être considéré comme un échantillon dont la variation de pondération est suffisamment modeste pour qu'il ne soit pas nécessaire de réduire l'effet de la variable de pondération sur les estimations, tel qu'une réduction de la pondération (voir Potter, 1988).

C.4 Effet du plan de sondage global sur la précision des estimations de l'enquête et sur l'homogénéité intragroupe dans les unités primaires d'échantillonnage (UPE) des principales estimations correspondantes de l'enquête

Environnement de travail

Kish (1965) a initialement défini un « effet du plan de sondage » noté $deff$ pour mesurer l'évolution multiplicatrice

de la variance d'une estimation de l'enquête due à l'utilisation de l'échantillonnage en grappes pour un plan d'échantillonnage donné, par rapport à la variance d'une estimation comparable provenant d'un échantillon aléatoire simple de la même taille. Chacun des résultats des principales mesures de l'enquête GATS est une proportion (p), ou un taux pouvant être exprimé sous la forme d'un pourcentage ou d'une moyenne. Le modèle original de Kish de l'effet du plan de sondage (dû à l'échantillonnage en grappes) d'une estimation

($\hat{p}$) de p est simplement

$$Deff_{CS} = 1 + \rho_{\theta}(\bar{m} - 1),$$

où ρ_{θ} est une mesure de l'homogénéité relative intragroupe des répondants au sein d'une UPE pour la mesure d'enquête utilisée en vue d'estimer p (mesurant le degré de comparabilité des membres de la même grappe d'échantillon avec l'ensemble des membres au sein de la population), $\bar{m} = n/a$ est le nombre moyen de répondants par UPE de l'échantillon, n est la taille de l'échantillon global des répondants, et a est le nombre d'UPE au sein de l'échantillon.

Quand un échantillon de l'enquête implique des grappes d'échantillonnage et une pondération inégale de l'échantillon, l'effet du plan de sondage global sur une estimation de l'enquête, désignée par $Deff(\hat{\theta})$, est le produit des deux dû à la fois à l'échantillonnage en grappes (voir ci-dessus) et à l'effet multiplicateur de la variable de pondération de l'échantillon (voir la discussion au sujet de $Meff_w$) comme démontré par Gabler *et al.* (1999), à savoir,

$$Deff(\hat{\theta}) \equiv \frac{V(\hat{p}; design, n)}{V(\hat{p}_{srs}; srs, n)}$$

$$= [Meff_w][Deff_{CS}(\hat{\theta})] = \left[1 + \frac{s_w^2}{\bar{w}^2}\right] [1 + \rho_\theta(\bar{m} - 1)] \quad (1)$$

Source(s) de données

Le fichier final de données pondérées utilisé pour l'analyse doit être utilisé pour ces calculs.

Logiciel de calcul

Les mêmes logiciels d'analyse d'enquête que ceux présentés dans le calcul de la marge d'erreur (voir l'annexe C.5).

Calcul

La variance estimée ou erreur-type (à savoir, $v(\hat{p}; design, n)$ ou $\sqrt{v(\hat{p}; design, n)}$ dans l'Équation 1, respectivement) est établie par le logiciel d'analyse d'enquête. Pour estimer ρ_θ , nous pouvons simplement résoudre ρ_θ dans l'Équation 1 et nous obtenons

$$\rho_\theta = \frac{\frac{Deff(\hat{\theta})}{Meff_w} - 1}{\bar{m} - 1} \quad (2)$$

Afin d'estimer ρ_θ pour chaque estimation de l'enquête, nous pouvons calculer $Meff_w$ à partir de l'échantillon. En effet, nous connaissons $\bar{m}$ car nous connaissons le nombre d'UPE (a) et la taille de l'échantillon global des répondants (n), et de nombreux logiciels d'analyse d'enquête (par exemple, SUDAAN) permettent de demander une estimation de $Deff(\hat{\theta})$ ainsi que l'estimation ($\hat{\theta}$). Si une estimation de $Deff(\hat{\theta})$ n'est pas utilisable directement à partir du produit du logiciel en déterminant $Deff(\hat{\theta})$ tel que défini dans l'Équation 1 ou ρ_θ comme le montre l'Équation 2, alors pour les proportions estimées, nous pouvons estimer l'effet de plan de sondage global comme suit :

$$deff(\hat{\theta}) = \frac{v(\hat{p}; design, n)}{v(\hat{p}_{srs}; srs, n)} = \frac{v(\hat{p}; design, n)}{\hat{p}(1 - \hat{p})/n},$$

puisque $v(\hat{p}_{srs}; srs, n) = \hat{p}(1 - \hat{p})/n$. L'estimation de ρ_θ peut alors être directement obtenue à partir de l'Équation 2.

Interprétation

On espère généralement que l'effet du plan de sondage global pour une estimation, ou $Deff(\hat{\theta})$, sera inférieur à 2,50. Les estimations de ρ_{θ} (de l'homogénéité au sein des UPE) sont généralement de petits nombres positifs compris entre 0 et 1 (le plus souvent entre 0 et 0,05). Dans certains cas, les estimations de ρ_{θ} seront de petits nombres négatifs.

Pour résumer les niveaux de $deff(\hat{\theta})$ parmi toutes les estimations de l'enquête GATS, il est recommandé d'estimer $Deff(\hat{\theta})$ pour plusieurs mesures principales des résultats de l'enquête (par exemple, la prévalence du tabagisme actuel, la prévalence du tabagisme ancien, le taux de sevrage, etc.) selon des domaines de rapports divers (par exemple, par genre, âge, zone urbaine/rurale, région, et diverses classifications croisées de ces variables). Ensuite, pour chaque résultat principal de mesure, résumez les valeurs estimées de $Deff(\hat{\theta})$ parmi tous les domaines de rapports, en rapportant les synthèses de statistiques suivantes pour l'estimation des valeurs de $Deff(\hat{\theta})$: médiane, minimum et maximum.

Résumez les valeurs de $\hat{\rho}_{\theta}$ (médiane, minimum et maximum) pour toutes les estimations de l'enquête GATS en procédant de la même manière que pour $deff(\hat{\theta})$.

Certains systèmes d'enquête ou organismes préfèrent rapporter l'effet de plan de sondage en le dénommant « deft » plutôt que « deff ». Deft est calculé pour chaque estimation comme le rapport entre l'erreur-type utilisant le plan d'échantillonnage donné et l'erreur-type que produirait l'utilisation d'un échantillon aléatoire simple.

C.5 Marge d'erreur pour les principales estimations de l'enquête

Environnement de travail

Commençons par rappeler quelques définitions élémentaires des mesures de la précision statistique des échantillons. En premier lieu, notez que toutes les estimations à partir d'échantillons sont susceptibles de différer légèrement de la situation réelle au sein de la population.

Cet écart est communément appelé l'**erreur d'échantillonnage**. Les échantillons aléatoires tels que ceux

utilisés dans le cadre de l'enquête GATS nous permettent d'établir des mesures sommaires de l'erreur d'échantillonnage, indiquant ainsi le degré de précision des estimations. Il existe plusieurs mesures sommaires du degré de précision des estimations de l'enquête. Afin d'établir ces mesures, nous utilisons le symbole θ pour désigner la caractéristique de la population que nous souhaitons estimer (par exemple le taux de prévalence des personnes qui fument actuellement des cigarettes). Nous utilisons le symbole $\hat{\theta}$ pour représenter l'estimation de θ à partir de l'échantillon.

La première mesure sommaire de la précision est la **variance**, que l'on écrit $V(\hat{\theta})$. Comme nous le verrons, toutes les autres mesures sont liées de près ou de loin à $V(\hat{\theta})$. La variance des estimations de l'enquête, et toutes les autres mesures connexes, sont des indicateurs quantitatifs de la variabilité des estimations de l'échantillon parmi tous les échantillons pouvant être produits par le plan d'échantillonnage

utilisé dans le cadre de l'enquête. Une autre mesure de la précision est l'**erreur-type** de l'estimation, définie comme $SE(\hat{\theta}) = \sqrt{V(\hat{\theta})}$. L'erreur-type d'une estimation est donc simplement la racine carrée de la variance de cette estimation ; elle est calculée dans les mêmes unités que l'estimation $\hat{\theta}$. Une troisième mesure de la précision est l'**erreur-type relative** de l'estimation, définie comme $RSE(\hat{\theta}) = SE(\hat{\theta}) / \theta = \sqrt{V(\hat{\theta})} / \theta$. Puisque $RSE(\hat{\theta})$ mesure la précision par rapport à la taille de ce qui est estimé, cette mesure n'a pas d'unité et constitue donc un meilleur indicateur de précision que $SE(\hat{\theta})$ pour comparer les estimations.

La mesure recommandée de la précision pour que les pays rendent compte des principales estimations de l'enquête GATS (par exemple, le taux de prévalence du tabagisme, etc.) est la **marge d'erreur**, définie comme $MOE(\hat{\theta}) = [Z][SE(\hat{\theta})]$, où Z est une mesure du niveau de confiance pour la mesure et $SE(\hat{\theta})$ est l'erreur-type de $\hat{\theta}$. Pour la plupart des estimations nationales de l'enquête GATS, nous pourrions utiliser $Z = 1.96$ pour un niveau de confiance de 95 %. La valeur estimée ($moe(\hat{\theta})$) rapportée pour $\hat{\theta}$ est interprétée comme suit :

Nous sommes sûrs à 95 % que la valeur déclarée ($\hat{\theta}$) se situe dans le champ de la valeur $moe(\hat{\theta})$ de θ .

Il est demandé aux analystes de l'enquête GATS d'indiquer la valeur de $moe(\hat{\theta})$ pour toutes les estimations principales, y compris toutes les estimations nationales relatives aux principaux indicateurs du tabagisme, ainsi que la valeur de $moe(\hat{\theta})$ pour les estimations régionales correspondantes si le plan d'échantillonnage exige des tailles d'échantillons régionaux accrues pour permettre la production d'estimations régionales qui répondront aux normes de précision de l'enquête GATS.

Étant donné que les mesures de la précision peuvent être aisément obtenues au moyen d'un logiciel statistique, les pays participant à l'enquête GATS sont vivement encouragés à régulièrement demander des estimations de la précision de toutes les estimations produites, même s'il n'est nécessaire de communiquer que celles relatives aux principales estimations nationales (et éventuellement régionales). Il est recommandé de communiquer les estimations pour lesquelles $MOE(\hat{\theta})$ avoisine ou dépasse la taille de l'échantillon lui-même avec prudence, voire de ne pas les communiquer. Il convient également de noter que tout comme $\hat{\theta}$, les estimations de $V(\hat{\theta})$, $SE(\hat{\theta})$, $MOE(\hat{\theta})$, ou $RSE(\hat{\theta})$ sont sujettes à des erreurs d'échantillonnage. Ceci est important, car toutes ces estimations peuvent être imprécises si elles sont fondées sur de petits échantillons, en particulier ceux comportant moins de 10 à 20 UPE.

Source(s) de données

Le fichier final de données pondérées utilisé pour l'analyse doit être utilisé pour ces calculs.

Logiciel de calcul

Les estimations des caractéristiques de la population ($\hat{\theta}$) et leurs valeurs correspondantes de $moe(\hat{\theta})$ doivent tenir compte des caractéristiques statistiquement significatives du plan d'échantillonnage. Les caractéristiques estimées doivent être pondérées, et les mesures de précision associées doivent permettre l'utilisation de la stratification, de l'échantillonnage en grappes, de l'échantillonnage sans remplacement et de la pondération de l'échantillon. Il est admis que le non-respect de cette recommandation entraîne généralement des estimations biaisées (en particulier de la précision) et donc des estimations défaillantes de l'intervalle et des tests de signification. Par conséquent, les analystes des pays participant à l'enquête GATS sont vivement encouragés à utiliser un logiciel d'analyse qui permette de pleinement tenir compte du plan d'échantillonnage utilisé pour produire les données de l'enquête. Cela signifie qu'il faut pondérer l'échantillon pour produire toutes les estimations (de $\hat{\theta}$ et $moe(\hat{\theta})$) destinées à l'analyse descriptive. Cela signifie également qu'il faut utiliser un logiciel qui applique une approche largement acceptée pour estimer les variances et les erreurs-types des estimations de l'enquête. Les trois principales méthodes d'estimation de la variance sont la linéarisation par série de Taylor (TSL), les répliques répétées équilibrées (BRR), et la méthode Jackknife (JRR) (Wolter, 1985). La plupart des logiciels d'analyse d'enquête utilisent la méthode TSL pour l'estimation de la variance. Pour des raisons de comparabilité entre les résultats des pays participant à l'enquête GATS, il est donc recommandé d'utiliser un logiciel utilisant l'approche TSL.

Plusieurs organismes statistiques à travers le monde ont développé des logiciels d'analyse de données tirées d'échantillons complexes tels que ceux de l'enquête GATS¹. Au-delà de la production d'estimations sur le tabagisme ($\hat{\theta}$), ces logiciels peuvent également générer des estimations de la précision (généralement $V(\hat{\theta})$ ou $SE(\hat{\theta})$) qui tiennent dûment compte des principales caractéristiques de conception de l'enquête GATS, à savoir l'échantillonnage en grappes, l'utilisation de la stratification, ainsi que la variation des probabilités de sélection (à savoir, la pondération de l'échantillon) :

Informations supplémentaires

Pour obtenir de plus amples détails au sujet des définitions fournies ci-dessus, nous vous invitons à consulter l'*Encyclopedia of Survey Research Methods* (2007). Pour accéder à des informations en ligne sur tous les logiciels d'analyse d'enquête, vous pouvez consulter le site internet financé par la Survey Research Methods Section of the American Statistical Association (<http://www.hcp.med.harvard.edu/statistics/survey-soft/>). Outre des liens vers les différents logiciels d'analyse, ce site fournit plusieurs autres liens et ressources utiles au sujet de la conception et de la mise en œuvre d'enquêtes (voir <http://www.amstat.org/sections/SRMS/index.html>).

¹ Le Centre de coordination des données de l'enquête GATS au sein du CDC fournit une assistance technique pour l'utilisation des logiciels de statistiques suivants : SAS, SPSS et STATA.

Calcul

Les produits des logiciels énumérés ci-dessus fourniront une valeur de $\hat{\theta}$ ainsi que de sa variance estimée, notée $v(\hat{\theta})$, ou de son erreur-type, notée $se(\hat{\theta})$. À partir de ces valeurs rapportées, on peut calculer la marge d'erreur estimée de $\hat{\theta}$ comme

$$moe(\hat{\theta}) = [t][se(\hat{\theta})] = [t]\sqrt{v(\hat{\theta})} \quad (1)$$

Interprétation

La valeur de $moe(\hat{\theta})$ rapportée pour $\hat{\theta}$ est interprétée comme suit :

Nous sommes sûrs à 95 % que la valeur estimée ($\hat{\theta}$) est comprise entre la valeur $moe(\hat{\theta})$ de .

C.6 Estimation des erreurs d'échantillonnage

Les estimations d'un échantillon d'enquête sont affectées par deux types d'erreur : 1) les erreurs non liées à l'échantillonnage et 2) les erreurs d'échantillonnage. Les *erreurs non liées à l'échantillonnage* sont le résultat d'erreurs ou de fautes qui ne peuvent être attribuables à l'échantillonnage et qui ont été commises lors de la réalisation de la collecte de données et du traitement des données, par exemple des erreurs de couverture, des erreurs de réponse, des erreurs de non-réponse, des questionnaires comportant des erreurs, des erreurs d'enregistrement de l'enquêteur, des erreurs de traitement de données, etc.

Malgré les nombreux efforts déployés lors de la mise en œuvre de l'enquête GATS ([pays]) afin de minimiser ces erreurs, les erreurs non liées à l'échantillonnage sont inévitables et difficiles à évaluer statistiquement.

L'échantillon des répondants sélectionnés pour l'enquête GATS ([pays]) n'est que l'un des échantillons qui pouvaient être sélectionnés à partir de la même population, en utilisant le même plan d'échantillonnage et la même taille d'échantillon. Chacun de ces échantillons pourrait produire des résultats susceptibles de différer légèrement des résultats de l'échantillon effectivement sélectionné. Les *erreurs d'échantillonnage* sont une mesure de la variabilité entre tous les échantillons possibles. L'étendue de la variabilité n'est pas connue exactement, mais peut être estimée statistiquement à partir des résultats de l'enquête.

Les mesures d'erreurs d'échantillonnage suivantes sont présentées pour chaque indicateur sélectionné :

- **Valeur (R)** : estimation de la prévalence pondérée de l'indicateur
- **Erreur-type (SE)** : Les erreurs d'échantillonnage sont généralement mesurées sous forme d'erreurs-types pour une estimation ou un indicateur (R) précis.
L'erreur-type d'une estimation est donc simplement la racine carrée de la variance de cette estimation ; elle est calculée dans les mêmes unités que l'estimation.
- **Taille de l'échantillon (n)** : nombre total d'observations utilisées pour calculer l'estimation de la prévalence (R).
- **Effet du plan de sondage** : L'*effet du plan de sondage*, noté « deff », est le rapport entre la variance réelle d'un indicateur – selon la méthode d'échantillonnage utilisée dans l'enquête – et

la variance calculée selon l'hypothèse d'un échantillonnage aléatoire simple. La racine carrée de l'effet du plan de sondage (« deft ») est utilisée pour indiquer l'efficacité du plan d'échantillonnage. Elle est calculée, pour chaque estimation, comme le rapport entre l'erreur-type utilisant le plan d'échantillonnage donné et l'erreur-type que produirait l'utilisation d'un échantillon aléatoire simple. Une valeur de DEFT de 1,0 indique que le plan d'échantillonnage est aussi efficace qu'un échantillon aléatoire simple. À l'inverse, une valeur de DEFT supérieure à 1,0 dénote une augmentation de l'erreur-type due à l'utilisation d'un plan d'échantillonnage plus complexe. En général, si l'enquête est bien conçue, DEFT varie entre 1 et 3. Il est courant, cependant, que DEFT soit beaucoup plus élevé, atteignant des valeurs de 7 ou 8.

- **Erreur-type relative (RSE)** : L'erreur-type relative, également appelée coefficient de variation (CV), est le rapport entre l'erreur-type et la valeur de l'indicateur.
- **Marge d'erreur (MOE)** : La marge d'erreur correspond à la mesure de la confiance désirée multipliée par l'erreur-type de l'estimation. Le niveau de confiance est généralement basé sur une valeur (Z) de la distribution normale standard. Par exemple, pour un niveau de confiance de 95 %, nous pouvons utiliser $Z = 1,96$.
- **Les limites de confiance** ($R \pm 1.96SE$) sont calculées pour indiquer l'intervalle dans lequel la vraie valeur pour la population peut raisonnablement se situer. Pour toute statistique calculée à partir de l'enquête, la valeur de celle-ci se situera dans une fourchette de plus ou moins deux fois l'erreur-type de cette statistique dans 95 % de tous les échantillons possibles de taille et de plan d'échantillonnage identiques.

Calcul de l'erreur-type

Si l'échantillon de répondants avait été choisi comme un échantillon aléatoire simple, il aurait été possible d'utiliser des formules simples pour calculer les erreurs d'échantillonnage. Toutefois, l'échantillon de l'enquête GATS de [année] est le résultat d'un plan d'échantillonnage stratifié en plusieurs étapes, qui a donc exigé l'utilisation de formules plus complexes. Le calcul des erreurs d'échantillonnage provenant des données de l'enquête GATS ([pays]) a été réalisé avec [version du logiciel de statistique]. La méthode utilisée pour déterminer les estimations de l'enquête correspondant à des moyennes ou à des pourcentages est la méthode de linéarisation de Taylor pour l'estimation de la variance.

La méthode de linéarisation de Taylor traite tout pourcentage ou moyenne comme une estimation d'un rapport $r = y/x$, où y représente la valeur totale de l'échantillon pour la variable y , et x représente le nombre total d'enquêtes dans le groupe ou sous-groupe étudié. La variance de r est calculée en appliquant la formule ci-dessous :

$$SE^2(r) = \text{var}(r) = \frac{1-f}{x^2} \sum_{h=1}^2 \left[\frac{m_h}{m_h - 1} \left(\sum_{i=1}^{m_h} Z_{hi}^2 - \frac{Z_h^2}{m_h} \right) \right]$$

dans laquelle $Z_{hi} = y_{hi} - rx_{hi}$, et $Z_h = y_h - rx_h$

où h (=1 ou 2) représente la strate urbaine ou rurale,
 m_h représente le nombre total d'UPE sélectionnées dans la h^{e} strate,
 y_{hi} est la somme des valeurs pondérées de la variable y dans la i^{e} UPE de la h^{e} strate,
 x_{hi} est la somme du nombre pondéré d'enquêtes dans la i^{e} UPE de la h^{e} strate, et

f est la fraction d'échantillonnage global, qui est ignorée car extrêmement faible.

Les résultats sont présentés dans cette annexe pour le pays dans son ensemble, pour les zones urbaines et rurales, ainsi que par genre. Pour chaque variable ou indicateur, le type de statistique (moyenne, pourcentage ou taux) ainsi que la population de base sont fournis dans le *tableau C-5*. Outre l'erreur-type décrite ci-dessus, le *tableau C-6* indique la valeur de l'estimation (R), la taille de l'échantillon, l'effet du plan de sondage (DEFF ou DEFT), l'erreur-type relative (SE/R), la marge d'erreur (MOE) ainsi que les limites de confiance à 95 % ($R \pm 1,96SE$) de chaque variable ou indicateur.

Tableau C-5 : Liste des indicateurs pour les erreurs d'échantillonnage, GATS [pays] [année]

Indicateur	Estimation	Population de référence
Consommateurs actuels de tabac	Pourcentage	Adultes âgés de 15 ans ou plus
Fumeurs actuels de tabac	Pourcentage	Adultes âgés de 15 ans ou plus
Fumeurs actuels de cigarettes	Pourcentage	Adultes âgés de 15 ans ou plus
Consommateurs actuels de tabac sans fumée	Pourcentage	Adultes âgés de 15 ans ou plus
Consommateurs quotidiens de tabac	Pourcentage	Adultes âgés de 15 ans ou plus
Fumeurs quotidiens de tabac	Pourcentage	Adultes âgés de 15 ans ou plus
Fumeurs quotidiens de cigarettes	Pourcentage	Adultes âgés de 15 ans ou plus
Consommateurs quotidiens de tabac sans fumée	Pourcentage	Adultes âgés de 15 ans ou plus
Anciens consommateurs quotidiens de tabac parmi tous les adultes	Pourcentage	Adultes âgés de 15 ans ou plus
Anciens consommateurs de tabac parmi les consommateurs quotidiens de tabac (actuels et anciens)	Pourcentage	Consommateurs quotidiens de tabac âgés de 15 ans ou plus (actuels et anciens)
Anciens fumeurs quotidiens de tabac parmi tous les adultes	Pourcentage	Adultes âgés de 15 ans ou plus
Anciens fumeurs de tabac parmi les fumeurs quotidiens (actuels et anciens)	Pourcentage	Fumeurs actuels ou anciens de tabac âgés de 15 ans ou plus
Anciens consommateurs quotidiens de tabac sans fumée parmi tous les adultes	Pourcentage	Adultes âgés de 15 ans ou plus
Anciens consommateurs quotidiens de tabac sans fumée parmi les consommateurs quotidiens de tabac sans fumée (actuels et anciens)	Pourcentage	Consommateurs quotidiens actuels et anciens de tabac sans fumée âgés de 15 ans ou plus
Première consommation de tabac dans les 5 minutes suivant le réveil	Pourcentage	Consommateurs quotidiens de tabac âgés de 15 ans ou plus
Première consommation de tabac dans les 6 à 30 minutes suivant le réveil	Pourcentage	Consommateurs quotidiens de tabac âgés de 15 ans ou plus
Tentative d'arrêt de fumer au cours des 12 derniers mois	Pourcentage	Fumeurs actuels ou anciens qui ont arrêté de fumer depuis moins de 12 mois
Tentative d'arrêt du tabac sans fumée au cours des 12 derniers mois	Pourcentage	Consommateurs actuels ou anciens de tabac sans fumée qui ont arrêté de consommer du tabac sans fumée depuis moins de 12 mois
Consultation d'un professionnel de santé au sujet du tabagisme	Pourcentage	Fumeurs actuels ou anciens qui ont arrêté de fumer depuis moins de 12 mois et qui ont consulté un professionnel de santé au cours des 12 derniers mois
Un professionnel de santé a conseillé d'arrêter de fumer	Pourcentage	Fumeurs actuels ou anciens qui ont arrêté de fumer depuis moins de 12 mois et qui ont consulté un professionnel de santé au cours des 12 derniers mois
Recours à la pharmacothérapie pour arrêter de fumer	Pourcentage	Fumeurs actuels ou anciens qui ont arrêté de fumer depuis moins de 12 mois
Recours à des services de conseil ou à des services téléphoniques d'aide à l'arrêt du tabac	Pourcentage	Fumeurs actuels ou anciens qui ont arrêté de fumer depuis moins de 12 mois
Prévoient, envisagent ou ont décidé d'arrêter	Pourcentage	Fumeurs actuels âgés de 15 ans ou plus
Exposition au tabagisme passif à la maison	Pourcentage	Adultes âgés de 15 ans ou plus
Exposition au tabagisme passif sur le lieu de travail	Pourcentage	Adultes travaillant en intérieur
Exposition au tabagisme passif dans les bâtiments/bureaux publics	Pourcentage	Adultes âgés de 15 ans ou plus

Global Adult Tobacco Survey (GATS)