

**ENQUÊTE SUR LA SANTÉ DANS LES COLLECTIVITÉS CANADIENNES /
ENQUÊTE SUR L'ACCÈS AUX SERVICES DE SANTÉ
(2003)
DOCUMENTATION DU FICHIER MAÎTRE**

STATISTIQUE CANADA

Décembre 2004

Table des matières

1.	INTRODUCTION	1
2.	CONTEXTE ET OBJECTIFS	2
3.	CONTENU DE L'ENQUÊTE	4
4.	PLAN D'ÉCHANTILLONNAGE	5
4.1	POPULATION CIBLE.....	5
4.2	BASE DE SONDAGE.....	5
4.3	TAILLE ET RÉPARTITION DE L'ÉCHANTILLON.....	5
4.4	RÉPARTITION DE L'ÉCHANTILLON INFRAPROVINCIAL.....	7
4.5	SÉLECTION DE L'ÉCHANTILLON DE LA PARTIE A.....	7
4.6	SÉLECTION DE L'ÉCHANTILLON DE LA PARTIE B.....	7
4.7	RÉPARTITION DE L'ÉCHANTILLON SUR LA PÉRIODE DE COLLECTE DES DONNÉES.....	8
5.	COLLECTE DES DONNÉES	9
5.1	DÉVELOPPEMENT DU QUESTIONNAIRE ET MÉTHODE DE COLLECTE DES DONNÉES.....	9
5.2	SUPERVISION ET CONTRÔLE.....	9
5.3	TECHNIQUES D'INTERVIEW.....	9
6.	TRAITEMENT DES DONNÉES	10
6.1	VÉRIFICATION.....	10
6.2	CODAGE.....	10
6.3	CRÉATION DE VARIABLES DÉRIVÉES ET GROUPÉES.....	10
6.4	PONDÉRATION.....	10
7.	PONDÉRATION	11
7.1	PROCÉDURES DE PONDÉRATION DE L'ESCC.....	11
7.2	PROCÉDURES DE PONDÉRATION DE L'EASS.....	12
7.2.1	ÉTAPES DE PONDÉRATION.....	13
8.	QUALITÉ DES DONNÉES	21
8.1	TAUX DE RÉPONSE.....	21
8.2	ERREURS DANS LES ENQUÊTES.....	23
8.2.1	ERREURS NON DUES À L'ÉCHANTILLONNAGE.....	23
8.2.2	ERREURS DUES À L'ÉCHANTILLONNAGE.....	24
8.2.2.1	MÉTHODE BOOTSTRAP POUR L'ESTIMATION DE LA VARIANCE.....	24
8.2.2.2	PROGRAMME BOOTVAR POUR L'ESTIMATION DE LA VARIANCE.....	25
9.	LIGNES DIRECTRICES POUR LA TOTALISATION, L'ANALYSE ET LA DIFFUSION	26
9.1	LIGNES DIRECTRICES POUR L'ARRONDISSEMENT.....	26
9.2	LIGNES DIRECTRICES POUR LA PONDÉRATION DE L'ÉCHANTILLON EN VUE DE LA TOTALISATION..	27
9.2.1	DÉFINITIONS DES CATÉGORIES D'ESTIMATIONS : DE TYPE NOMINAL PAR OPPOSITION À QUANTITATIVES.....	27
9.2.2	TOTALISATION D'ESTIMATIONS DE TYPE NOMINAL.....	28
9.2.3	TOTALISATION D'ESTIMATIONS QUANTITATIVES.....	28
9.3	LIGNES DIRECTRICES POUR L'ANALYSE STATISTIQUE.....	28
9.4	LIGNES DIRECTRICES POUR LA DIFFUSION.....	29

10.	UTILISATION DU FICHIER	31
10.1	UTILISATION DES FACTEURS DE PONDÉRATION	31
10.2	CONVENTION APPLIQUÉE POUR NOMMER LES VARIABLES	31
10.2.1	STRUCTURE ÉLÉMENTAIRE DES NOMS DES VARIABLES POUR L'EASS.....	31
10.2.2	POSITIONS 1 À 3 : NOM DE LA VARIABLE/SECTION DU QUESTIONNAIRE	32
10.2.3	POSITION 4 : CYCLE.....	32
10.2.4	POSITION 5 : TYPE DE VARIABLE	33
10.2.5	POSITIONS 6 À 8 : NOM DE LA VARIABLE	33
10.3	ACCÈS AUX FICHIERS MAÎTRES	33

1. Introduction

L'Enquête sur l'accès aux services de santé (EASS) fut rendue publique pour la première fois en 2002 en tant que supplément à l'Enquête sur la santé dans les collectivités canadiennes (ESCC) 2000-2001, cycle 1.1, comme une enquête ayant lieu seulement une fois. Lors du cycle 2.1 de l'ESCC, il fut décidé d'incorporer l'EASS dans son contenu, mais de la rendre publique en tant que fichier séparé.

Ce fichier contient les données recueillies lors du deuxième cycle de collecte de l'EASS. Les informations ont été recueillies entre janvier 2003 et décembre 2003 pour les 126 régions sociosanitaires, dans les 10 provinces. L'EASS recueille les réponses des personnes âgées de 15 ans et plus, vivant dans des logements privés. Sont exclues de l'échantillon les individus vivant sur les réserves indiennes, les terres de la Couronne, les personnes vivant en institutions, les membres à temps plein des forces canadiennes, et les personnes vivant dans certaines régions éloignées.

Le présent document a pour but de faciliter la manipulation du fichier maître de l'EASS 2003, qui est décrit en détails dans le texte qui suit.

Pour toute question concernant les ensembles de données ou leur utilisation, s'adresser à :

Unité d'accès aux données, Division de la statistique de la santé:
Courriel :

1 (613) 951-1653
cchs-escc@statcan.ca

Totalisations spéciales ou renseignements généraux sur les données :

Services à la clientèle, Division de la statistique de la santé :
Courriel :

1 (613) 951-1746
hd-ds@statcan.ca

2. Contexte et objectifs

Suite à l'Accord des Premiers Ministres sur la santé en septembre 2000, les gouvernements fédéral, provinciaux et territoriaux se sont entendus pour rendre compte des résultats et de la performance des systèmes de santé aux citoyens à partir de septembre 2002. Selon cet accord, ces comptes-rendus devaient contenir des indicateurs de performance dans 14 domaines spécifiques à la santé. Dans le domaine de la qualité des services, les données pour deux de ces indicateurs étaient inexistantes à savoir : a) le temps d'attente pour certains tests diagnostiques et pour des traitements et b) l'accès en tout temps (24/7) aux services de santé. L'enquête sur l'accès aux services de santé a été mise sur pied afin de recueillir de telles informations ainsi que des renseignements concernant les expériences vécues par les patients, l'acceptabilité et les perceptions face à l'attente pour des soins.

Les informations concernant l'accès aux services de première ligne 24/7 comprennent :

- l'expérience des répondants dans l'obtention d'information ou de conseils de santé;
- l'expérience des répondants dans l'obtention de services de soins de santé de routine ou continus;
- l'expérience des répondants dans l'obtention de soins immédiats pour un problème de santé mineur comme la fièvre, des maux de tête, une foulure, des vomissements, des irruptions cutanées, etc.;
- l'expérience des répondants dans l'obtention des services de santé en général.

On demandait aux répondants leur avis quant à l'obtention de services de première ligne à différentes heures de la journée, les problèmes rencontrés pour obtenir des services et l'endroit où ils ont obtenu les services.

Les informations concernant l'accès aux services spécialisés comprennent :

- l'expérience des répondants ayant besoin d'un spécialiste tel qu'un cardiologue, un allergiste, etc. en vue d'obtenir un diagnostic pour un nouveau problème de santé;
- l'expérience des répondants ayant besoin d'une chirurgie non urgente telle qu'une chirurgie cardiaque, une chirurgie aux articulations, etc.;
- l'expérience des répondants ayant besoin de certains tests diagnostiques tels que IRM, tomographie, angiographies.

On demandait aux répondants leur avis quand à l'accès à ces services, le temps d'attente, l'acceptabilité et les conséquences de ce temps d'attente.

Cette enquête avait été réalisée en 2001 en tant que supplément à l'ESCC, comme étant une enquête effectuée une seule fois. En 2003, l'enquête a été amalgamée avec l'Enquête sur la santé dans les collectivités canadiennes (ESCC) cycle 2.1 et fait maintenant partie de ce programme statistique, malgré qu'elle soit parue en tant que fichier séparé. Lors du prochain cycle de l'ESCC, à paraître en 2006, cette enquête sera complètement amalgamée et il n'y aura désormais ni de fichier séparé ni de questionnaire séparé.

Le questionnaire de l'Enquête sur l'accès aux services de santé (EASS) comprend deux modules, l'un sur l'accès aux services de soins santé l'autre sur le temps d'attente. La collecte des données de l'EASS a été intégrée à celle de l'ESCC de la façon suivante :

Partie A : En tout, 18 981 personnes ont été interviewées pour l'EASS au moment de l'interview de l'ESCC. Tous ces répondants font aussi partie du sous-échantillon 3. (Les données sur le temps d'attente ne sont pas incluses dans le fichier de données de l'ESCC.) Ces interviews ont été réalisées de septembre à décembre 2003 (en outre, 45 interviews ont été réalisées en janvier 2004).

Partie B : Les intervieweurs ont repris contact avec, en tout, 13 024 personnes interviewées antérieurement pour l'ESCC. Ces répondants **ne faisaient pas** partie du sous-échantillon 3 lors de la première interview. On leur a posé les questions des modules sur l'accès aux services de santé et sur les temps d'attente lors de la deuxième interview. Les personnes qui ont répondu à la partie B de l'EASS n'ont pas été incluses dans le sous-échantillon 3.

3. Contenu de l'enquête

L'EASS a été conçue pour recueillir des données sur l'accès aux services de premier contact et aux services spécialisés. Les services de premier contact comprennent les soins de santé de routine, de l'information et des conseils en matière de santé, des soins immédiats pour un problème mineur de santé, prodigués par un médecin de famille ou un omnipraticien, une infirmière ou un autre fournisseur de soins, excluant les spécialistes. Les services spécialisés sont les visites chez un médecin spécialiste pour un diagnostic ou un nouveau problème de santé, une chirurgie non urgente et des tests diagnostiques choisis (IRM non urgente, tomodensitométrie et angiographie).

L'EASS 2003 comprenait deux modules de questionnaires : Accès aux services de santé et Temps d'attente. Dans le premier module, les questions étaient axées sur les difficultés d'accéder aux services de premier contact, par exemple les soins de routine, l'information et les conseils en matière de santé, ainsi que des soins immédiats pour traiter un problème de santé mineur, que ce soit pour la personne elle-même ou pour un membre de sa famille. On a demandé aux personnes qui avaient eu des difficultés de préciser les types de barrières auxquelles elles s'étaient heurtées. On leur a aussi demandé si elles avaient un médecin de famille régulier.

Le deuxième module de l'EASS portait sur l'accès aux services spécialisés tels que les visites chez un spécialiste pour un nouveau problème de santé ou condition, les interventions chirurgicales non urgentes et certains tests diagnostiques. Les temps d'attente constituent toujours un enjeu pour ce qui est de l'accès aux services spécialisés. L'EASS fournit des données comparables sur le temps d'attente déclaré par les répondants aux niveaux national et provincial. Afin d'assurer la comparabilité des données, on a eu recours à des définitions types de temps d'attente en vue d'obtenir une visite chez un spécialiste, de chirurgies non urgentes et de tests diagnostiques. On a demandé aux répondants s'ils estimaient que le temps qu'ils avaient attendu était acceptable et si le fait d'attendre avant d'obtenir des services spécialisés avait eu des conséquences sur leur vie. L'EASS fournit aussi des renseignements sur d'autres difficultés et barrières auxquels se heurtent les personnes au moment d'avoir accès aux services spécialisés.

4. Plan d'échantillonnage

4.1 Population cible

L'EASS vise la population de 15 ans et plus vivant à domicile des dix provinces canadiennes. Sont exclues du champ de l'enquête les personnes vivant dans les trois territoires, dans les réserves indiennes et les terres de la couronne, les personnes vivant en établissements, les membres à temps plein des Forces canadiennes que les personnes vivant dans certaines régions éloignées. L'EASS couvre environ 98 % de la population de 15 ans et plus vivant dans les dix provinces.

4.2 Base de sondage

Étant donné que l'EASS est un sous-échantillon de l'ESCC 2.1, l'EASS utilise les mêmes bases de sondage que l'ESCC, qui utilise plusieurs bases. D'abord, elle utilise la base aréolaire de l'Enquête sur la population active (EPA). Le plan d'échantillonnage de l'EPA est un plan d'échantillonnage en grappes stratifié à plusieurs degrés où le logement représente l'unité finale d'échantillonnage. L'ESCC utilise également deux types de bases téléphoniques : une base liste de numéros de téléphone et une base à composition aléatoire (CA). Pour plus de détails concernant ces bases de sondage, veuillez consulter le guide de l'utilisateur de l'ESCC 2.1¹.

4.3 Taille et répartition de l'échantillon

Afin de produire des estimations fiables pour les dix provinces, notamment en ce qui concerne les estimations des temps d'attente, les tailles d'échantillon suivantes (tableau 4.1) ont été établies.

Tableau 4.1 Tailles d'échantillon visées par province pour l'EASS

Province	Taille totale de l'échantillon (visée)
Terre-Neuve-et-Labrador	2 400
Île-du-Prince-Édouard	1 200
Nouvelle-Écosse	3 000
Nouveau-Brunswick	3 000
Québec	4 000
Ontario	4 800
Manitoba	3 200
Saskatchewan	3 200
Alberta	3 600
Colombie-Britannique	4 000
Canada	32 400

¹ Statistique Canada (2004). *ESCC Cycle 2.1, Fichier de microdonnées à grande diffusion*. Numéro de catalogue 82M0013XCB.

Pour atteindre les tailles d'échantillon visées, un sous-échantillon de répondants de l'ESCC a été choisi. Cependant, compte tenu de la date de la mise en place de l'EASS (été-automne 2003), la sélection du sous-échantillon de répondants de l'ESCC a dû se faire en deux parties. D'abord, une première partie de l'échantillon (appelée partie A) provient directement de l'ESCC. À partir de septembre 2003, on a ajouté le module sur les temps d'attente au questionnaire de l'ESCC et, celui-ci et les autres modules de l'EASS ont été posés à même l'interview de l'ESCC pour un sous-échantillon de l'ESCC entre les mois de septembre et décembre 2003. Cependant, pour certaines provinces, la taille d'échantillon de l'ESCC durant cette période était insuffisante pour atteindre les tailles d'échantillon visées pour l'EASS, même en sélectionnant 100 % des répondants de l'ESCC durant cette période. Pour combler les besoins en échantillon, il a fallu recontacter par téléphone un sous-échantillon des personnes qui avaient déjà répondu à l'ESCC entre janvier et septembre 2003. Cette partie de l'échantillon est appelée la partie B. Il est important de mentionner que par souci d'uniformité, la ventilation partie A/partie B de l'échantillon de l'EASS s'est avérée être sensiblement la même pour chacune des provinces même si ce n'était pas vraiment nécessaire dans certaines provinces. Le tableau 4.2 donne les tailles visées pour chacune des deux parties de l'EASS.

Étant donné que les répondants de l'ESCC étaient sélectionnés à partir de deux bases (la base aréolaire et les bases téléphoniques) nous avons également contrôlé la provenance de l'échantillon selon les bases de sondage pour l'échantillon de l'EASS. Pour la partie A, une répartition égale a été faite entre les bases de sondage car près de 50 % de l'échantillon de l'ESCC provient de la base aréolaire et l'autre 50 % des bases téléphoniques. Pour la partie B, puisque le recontact devait se faire par téléphone, nous avons alloué plus d'échantillon aux répondants de l'ESCC provenant de la base téléphonique. Mais, puisque dans trois provinces on devait recontacter tous les répondants de l'ESCC (éligibles) pour atteindre les tailles d'échantillon visées, et que ces répondants provenaient des deux bases de sondage, on a alloué de l'échantillon dans les deux types de base de sondage dans toutes les provinces par souci d'uniformité. Le tableau 4.2 donne également la répartition de l'échantillon par type de base de sondage.

Tableau 4.2 Tailles d'échantillon visées selon la partie de l'échantillon (A et B)

Province	Partie A (tailles visées)			Partie B (tailles visées)		
	Total	Base aréolaire	Base Tél.	Total	Base aréolaire	Base Tél.
Terre-Neuve-et-Labrador	880	440	440	1 520	720	800
Île-du-Prince-Édouard	560	280	280	640	310	330
Nouvelle-Écosse	1 200	600	600	1 800	850	950
Nouveau-Brunswick	1 160	580	580	1 840	740	1 100
Québec	3 000	1 500	1 500	1 000	250	750
Ontario	3 800	1 900	1 900	1 000	250	750
Manitoba	1 640	820	820	1 560	520	1 040
Saskatchewan	1 810	905	905	1 390	410	980
Alberta	3 000	1 500	1 500	600	200	400
Colombie-Britannique	3 200	1 600	1 600	800	200	600
Canada	20 250	10 125	10 125	12 150	4 450	7 700

4.4 Répartition de l'échantillon infraprovincial

L'ESCC vise à produire des estimations fiables à l'échelle des régions sociosanitaires tandis que l'EASS vise à produire des estimations fiables à l'échelle provinciale. Pour cette raison, la répartition de l'échantillon de l'ESCC par région sociosanitaire n'est pas optimale pour l'EASS. Dans les provinces où cela était possible, on a tenté de répartir l'échantillon de l'EASS à l'intérieur de la province de façon à ce que la taille d'échantillon de chaque région sociosanitaire se rapproche le plus possible de la taille d'échantillon obtenue à l'aide d'une répartition proportionnelle à la taille de la région (répartition optimale). Par exemple, la région de Montréal-Centre représente 25 % de la population du Québec. Cependant, dans l'ESCC, seulement 13 % de l'échantillon total du Québec est alloué à cette région. Par contre, pour l'EASS, nous avons alloué l'échantillon de sorte à ce que la région de Montréal-Centre reçoive 25 % de la taille d'échantillon total du Québec.

4.5 Sélection de l'échantillon de la partie A

Une fois que l'échantillon était alloué aux deux parties de l'EASS, l'échantillon de la partie A était sélectionné dans chaque région sociosanitaire par base de sondage (aréolaire et téléphone). La sélection de l'échantillon de la partie A consistait à choisir aléatoirement un sous-échantillon des logements déjà sélectionnés dans l'échantillon de l'ESCC (entre septembre et décembre 2003) selon la taille visée pour toutes les combinaisons « région / base de sondage ». Dans certaines provinces, l'échantillon de la partie A correspondait à tous les logements sélectionnés dans l'ESCC durant cette période. Notez que les tailles d'échantillon visées pour la partie A avaient été préalablement gonflées pour tenir compte de la non-réponse anticipée et des logements vacants.

4.6 Sélection de l'échantillon de la partie B

La partie B de l'EASS consistait à recontacter par téléphone un sous-échantillon de répondants de l'ESCC. Cependant, ce ne sont pas tous les répondants de l'ESCC qui étaient éligibles à être recontactés pour l'EASS. Les critères d'éligibilité étaient les suivants :

- i) Le répondant devait avoir été interviewé pendant des périodes de collecte de l'ESCC comprises entre janvier et août 2003 (inclusivement).
- ii) Le répondant devait avoir 15 ans ou plus lors de l'interview de l'ESCC.
- iii) Le répondant ne devait pas avoir fait partie du sous-échantillon 3 de l'ESCC (i.e, ne pas avoir déjà répondu au module « Accès aux services de soins de santé »).
- iv) L'interview de l'ESCC ne devait pas avoir été faite par procuration.
- v) Un numéro de téléphone valide et unique devait avoir été fourni par le répondant (ce critère s'appliquait seulement aux répondants provenant de la base aréolaire).

Une fois les répondants éligibles identifiés, la sélection de l'échantillon de la partie B consistait seulement à choisir aléatoirement un sous-échantillon des répondants éligibles selon la taille visée pour toutes les combinaisons « région / base de sondage ». Les tailles d'échantillon visées pour la partie B avaient été préalablement gonflées pour tenir compte de la non-réponse anticipée. Un taux de réponse de 80 % avait été prévu pour la partie B de l'EASS.

4.7 Répartition de l'échantillon sur la période de collecte des données

La collecte de la partie A a débuté au mois de septembre 2003 et s'est terminée en décembre 2003 (il y a eu quelques cas faits au début janvier 2004). L'échantillon de la partie A a été réparti selon la répartition de l'échantillon de l'ESCC, soit en trois périodes de collecte : septembre, octobre et novembre. En général, les tailles d'échantillon du mois de septembre et octobre étaient identiques tandis que la taille d'échantillon du mois de novembre a été la moitié de celles des mois de septembre et octobre.

Pour la partie B, il y avait une seule période de collecte. La collecte a débuté en octobre 2003 et s'est terminée en décembre 2003.

5. Collecte des données

5.1 Développement du questionnaire et méthode de collecte des données

Le questionnaire de l'EASS 2003 a fait l'objet d'interviews assistées par ordinateur (IAO). Des unités d'échantillonnage sélectionnées à partir de la base aréolaire ont répondu aux questions suivant la méthode d'interview en personne assistée par ordinateur (IPAO) tandis que les autres unités, sélectionnées à partir des bases de sondage téléphoniques, ont répondu aux questions suivant la méthode de l'interview téléphonique assistée par ordinateur (ITAO).

L'IAO procure un certain nombre d'avantages quant à la qualité des données par rapport aux autres méthodes de collecte. Premièrement, le libellé des questions, comprenant les périodes de référence et les pronoms, est personnalisé automatiquement en fonction de facteurs comme l'âge et le sexe du répondant, de la date de l'interview et des réponses aux questions précédentes. En second lieu, on applique des mesures de contrôle qui isolent les réponses incohérentes ou hors normes, et des prompts apparaissent à l'écran lorsqu'une entrée incorrecte est enregistrée. Le répondant reçoit une rétroaction immédiate et l'intervieweur peut corriger toute incohérence. Troisièmement, le processus permet de sauter automatiquement les questions qui ne concernent pas le répondant.

5.2 Supervision et contrôle

Les intervieweurs IPAO travaillaient séparément depuis leur domicile à l'aide d'un ordinateur portable et étaient supervisés à distance par des intervieweurs principaux. Les interviews complétées étaient envoyées quotidiennement de leur domicile au bureau central de Statistique Canada, et ce, par transmission téléphonique protégée.

Les intervieweurs ITAO travaillaient dans des bureaux centralisés sous la supervision d'un intervieweur principal. Il incombait au surveillant de projet du bureau régional, à l'intervieweur principal et à l'équipe d'assistance technique de transmettre les dossiers de chacun des cinq bureaux ITAO au bureau central.

On ne pouvait compter, pour la collecte par ITAO, sur un ordonnanceur d'appels automatique, c'est-à-dire un système central qui optimise l'agencement des rappels et l'ordonnancement des rendez-vous. Au lieu, un lot de cas était distribué au début de chaque période de collecte à chaque ordinateur de tous les bureaux ITAO. Le nombre de cas attribués à chaque ordinateur faisait alors l'objet d'un traitement manuel. À cause du nombre relativement restreint de cas réservés à l'ITAO, cette démarche s'est révélée raisonnablement efficace et l'absence d'un ordonnanceur d'appels ne semble pas avoir eu d'effet nuisible sur la qualité des données.

5.3 Techniques d'interview

Pour la partie A de l'EASS les intervieweurs IPAO étaient formés pour effectuer un premier contact personnel avec chaque personne sélectionnée. Dans les cas où la première visite entraînait une non-réponse, il était permis de faire un suivi par téléphone.

Toutes les entrevues de la partie B ont été faites au téléphone, soit par des intervieweurs IPAO travaillant de leur foyer ou par des intervieweurs travaillant dans des centres d'appels centralisés.

6. Traitement des données

6.1 Vérification

L'application IAO a effectué la plus grande partie de la vérification des données au moment de l'interview. Les intervieweurs ne pouvaient pas taper de valeurs hors-normes et les erreurs d'enchaînement faisaient l'objet de l'instruction de contrôle programmée « passez à ». Par exemple, l'IAO s'assurait de ne pas poser au répondant les questions non pertinentes.

En réponse à certaines données incompatibles ou inhabituelles, on a signalé des messages d'avertissement, mais sans prendre de mesures correctrices au moment de l'interview. On a plutôt mis au point, le cas échéant, des versions révisées à appliquer après la collecte des données au bureau central. Les incohérences ont été le plus souvent corrigées en attribuant à l'une ou aux deux variables en question la valeur « non déclaré ».

6.2 Codage

On a fourni des catégories de réponses précodées pour toutes les variables appropriées. Il est possible de donner une réponse ouverte à plusieurs questions du questionnaire de l'EASS 2003. Les réponses ouvertes à certaines de ces questions étaient codées à l'intérieur de l'une des catégories inscrites sur la liste à partir du moment où l'information écrite renvoyait à une catégorie sur la liste.

6.3 Création de variables dérivées et groupées

Pour faciliter l'analyse des données, on a dérivé un certain nombre de variables à partir des éléments trouvés dans le questionnaire de l'EASS 2003. Le cinquième caractère du nom des variables dérivées est en général un « D » ou un « G ». Dans certains cas, les variables dérivées sont simples, donnant lieu à un regroupement des catégories de réponses. Dans d'autres cas, on a combiné plusieurs variables pour en créer une nouvelle. Veuillez consulter la documentation sur les variables dérivées pour plus d'information.

6.4 Pondération

Le principe de base de l'estimation dans un échantillon aléatoire comme celui de l'EASS 2003 repose sur le fait que chaque personne représente, en plus d'elle-même, plusieurs autres personnes qui ne font pas partie de l'échantillon. Par exemple, dans un simple échantillon aléatoire de 2 % de la population, chaque personne en représente 50. Dans la terminologie en usage ici, nous dirons que nous avons attribué à chaque personne un facteur de pondération de 50.

L'étape de détermination des facteurs de pondération donne lieu au calcul du poids d'échantillonnage de chaque personne échantillonnée. Ce poids apparaît dans le fichier de microdonnées et doit servir à extraire des estimations de l'enquête. Par exemple, si l'on doit évaluer le nombre de personnes qui fument tous les jours, on le fait en choisissant dans l'échantillon les enregistrements des personnes qui présentent cette caractéristique et en faisant la somme des facteurs de pondération que représentent ces enregistrements. Vous trouverez les détails sur la façon dont on calcule les poids d'échantillonnage à la section 7.

7. Pondération

Le poids d'échantillonnage associé à un répondant de l'enquête peut être vu comme étant le nombre de personnes dans la population représentées par le répondant. Lors du calcul d'estimations ponctuelles, les poids d'échantillonnage doivent être utilisés avec les données d'enquête afin d'obtenir des estimations représentatives de la population couverte par l'enquête.

Tel que décrit à la section 5, l'EASS est un sous-échantillon de l'ESCC 2.1. Alors, la création des poids pour l'EASS est fortement liée aux procédures de pondération de l'ESCC. En fait, le poids de départ utilisé pour l'EASS est un des sous-poids intermédiaires de l'ESCC (après une certaine étape). Les procédures de pondération de l'ESCC sont décrites brièvement à la section 7.1; des explications détaillées des procédures de pondération de l'EASS sont décrites à la section 7.2.

7.1 Procédures de pondération de l'ESCC

À l'intérieur des 10 provinces, l'ESCC 2.1 a été planifiée pour produire des estimations fiables pour 126 régions sociosanitaires. Par conséquent, les bases de sondage utilisées pour l'ESCC ont été stratifiées selon ces 126 régions sociosanitaires. L'ESCC a utilisé deux types de bases de sondage; une base aréolaire et une base téléphonique. La taille d'échantillon visée de 130 000 répondants a été répartie entre les provinces et ensuite, entre les régions sociosanitaires à l'intérieur de chaque province. Finalement, à l'intérieur de chaque région sociosanitaire, l'échantillon alloué a été réparti entre les deux bases de sondage de façon à ce qu'environ la moitié de l'échantillon soit attribuée à chaque base de sondage. Noter que pour un petit nombre de régions sociosanitaires, seulement une base de sondage a été utilisée. La pondération de l'ESCC a été faite indépendamment pour les deux types de base de sondage. Une intégration des deux bases a par la suite été faite avant d'ajuster les poids pour la saisonnalité et d'effectuer une post-stratification. Les procédures de pondération de l'ESCC sont résumées dans le diagramme A ci-dessous.

Diagramme A **Sommaire de la stratégie de pondération de l'ESCC**

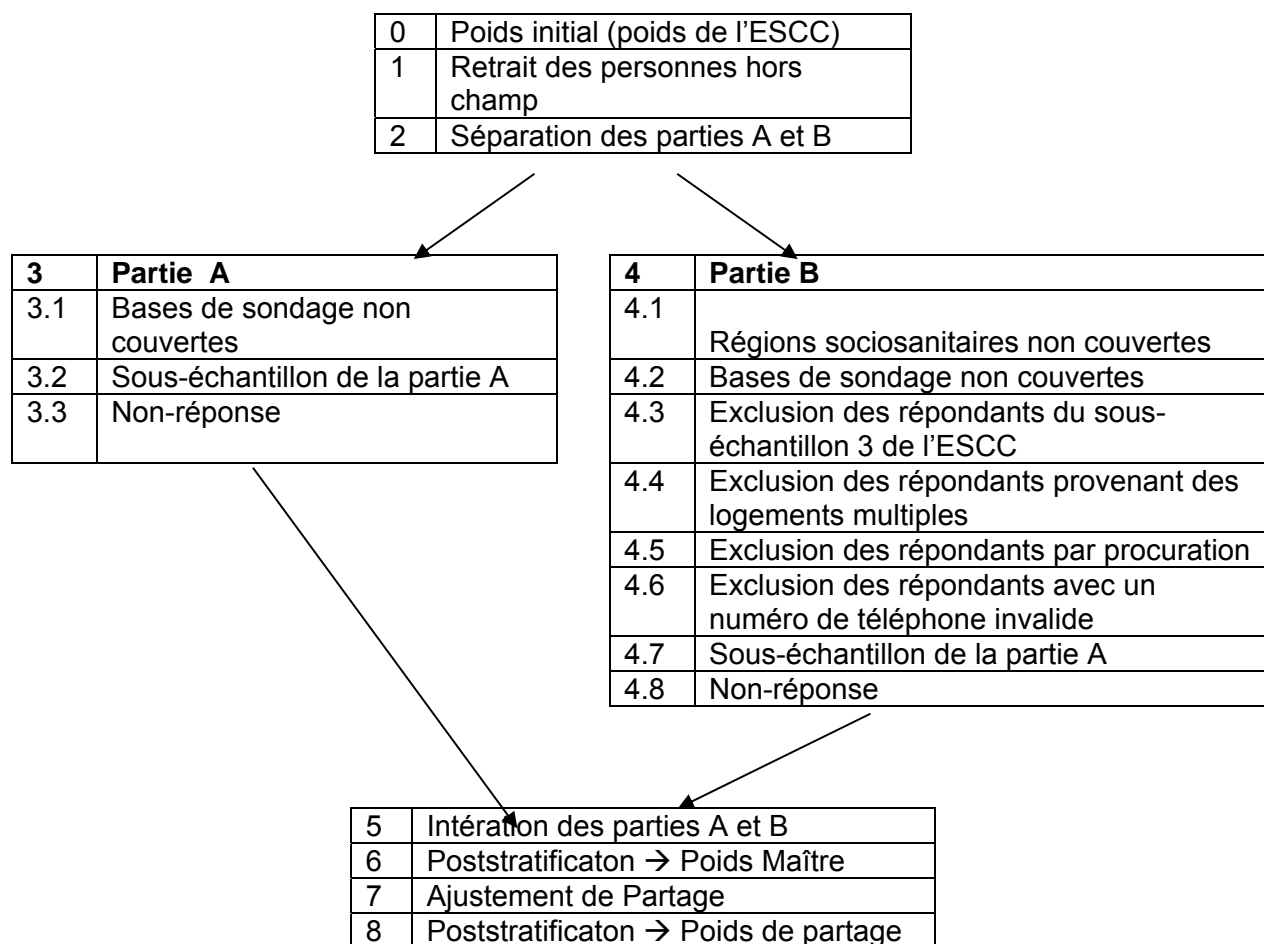
Base aréolaire	Base téléphonique
A0 - Poids initial	T0 - Poids initial
A1 - Accroissement de l'échantillon	T1 - Nombre de mois
A2 - Stabilisation	T2 - Retrait des unités hors champ
A3 - Retrait des unités hors champ	T3 - Couverture des bases listes
A4 - Non-réponse ménage	T4 - Combinaison des bases listes
A5 - Création du poids-personne	T5 - Non-réponse ménage
A6 - Non-réponse personne	T6 - Ménages sans téléphone
Poids final de la base aréolaire	T7 - Création du poids-personne
↘	T8 - Non-réponse personne
	T9 - Lignes multiples
	Poids final de la base téléphonique
	↙
	I1 - Intégration
	I2 - Effet saisonnier
	I3 - Poststratification
	Poids final du cycle 2.1 de l'ESCC

Pour obtenir plus de détails au sujet de la pondération de l'ESCC, veuillez consulter le guide de l'utilisateur de l'ESCC 2.1.

7.2 Procédures de pondération de l'EASS

Tel que décrit à la section 4, l'échantillon de l'EASS a été divisé en deux parties : la partie A et la partie B. Les poids de l'ESCC après l'intégration ont été utilisés comme poids initial pour la pondération des deux parties de l'EASS. La pondération de chacune des deux parties a été faite de façon indépendante et, à la fin, les poids des deux parties ont été intégrés. Une poststratification a par la suite été appliquée aux poids intégrés de l'EASS. Les procédures de pondération de l'EASS sont résumées dans le diagramme B ci-dessous.

Diagramme B Étapes de pondération de l'EASS



7.2.1 Étapes de pondération

Ajustement 0 – Poids initial

Les poids de l'ESCC après intégration ont été utilisés comme poids initiaux pour l'EASS. Ces poids sont appelés les poids 0 à l'EASS.

Ajustement 1 – Retrait des personnes hors champ

Contrairement à l'ESCC, la population cible de l'EASS ne comprend pas les groupes suivants : les personnes âgées de moins de 15 ans, les personnes vivant dans les trois territoires et les personnes habitant dans la région sociosanitaire des Terres-Cries-de-la-Baie-James. Les répondants de l'ESCC appartenant à un de ces groupes ont été considérés hors champ et enlevés de l'échantillon. Les répondants restant dans le champ de l'enquête ont gardé le même poids qu'à l'étape précédente, et ce poids est maintenant appelé le poids 1.

Ajustement 2 – Séparation des parties A et B

L'échantillon de l'EASS a été choisi en deux parties : la partie A et la partie B. La partie A consistait à sous-échantillonner les logements choisis dans l'ESCC lors des périodes de collecte de septembre à novembre 2003. La partie B consistait à sous-échantillonner les répondants de l'ESCC provenant des périodes de collecte de janvier à août 2003. Puisque pour ces deux parties, les stratégies de collecte étaient différentes et indépendantes, il a été décidé d'ajuster séparément les poids des répondants de l'ESCC associés à chacune des deux parties de façon à ce que chaque partie représente la population couverte par l'enquête (comme si c'était deux enquêtes indépendantes).

La sélection des sous-échantillons pour les parties A et B a été faite en contrôlant selon la base de sondage par région sociosanitaire. Tous les ajustements de poids et ce, jusqu'à l'ajustement pour la sélection des sous-échantillons, ont donc été faits par base de sondage et région sociosanitaire. Le facteur d'ajustement à cette étape a été calculé de la façon suivante :

Partie A

Pour les répondants de l'ESCC des périodes de collecte de septembre à novembre 2003 (éligibles pour la partie A de l'EASS), le facteur d'ajustement était le suivant :

$$\frac{\text{Somme des poids 1 pour tous les répondants de l' ESCC}}{\text{Somme des poids 1 pour tous les répondants de l' ESCC lors des périodes de collecte de sept. à nov. 2003}}$$

Le poids 1 des répondants de l'ESCC des périodes de collecte de septembre à novembre 2003 a été multiplié par ce facteur d'ajustement pour produire le poids 2.

Partie B

Pour les répondants de l'ESCC des périodes de collecte de janvier à août 2003 (éligibles pour la partie B de l'EASS), le facteur d'ajustement était le suivant :

$$\frac{\text{Somme des poids 1 pour tous les répondants de l' ESCC}}{\text{Somme des poids 1 pour tous les répondants de l' ESCC lors des périodes de collecte de jan. à août 2003}}$$

Le poids 1 des répondants de l'ESCC des périodes de collecte de janvier à août 2003 a été multiplié par ce facteur d'ajustement pour produire le poids 2.

Ajustement 3 – Étapes d'ajustement pour la partie A

En ce qui concerne les ajustements s'appliquant seulement à la partie A de l'EASS, seuls les répondants de l'ESCC lors des périodes de collecte de septembre à novembre 2003 ont été conservés. Pour ceux-ci, le poids 2 est maintenant appelé le poids 3. Les ajustements 3.1 à 3.3 ci-dessous s'appliquent seulement à la partie A de l'EASS.

Ajustement 3.1 – Bases de sondage non couvertes

Au départ, tous les répondants de l'ESCC des périodes de collecte de septembre à novembre 2003 étaient supposés avoir une chance d'être sélectionnés dans le sous-échantillon de la partie A de l'EASS. Des changements dans l'étape de sélection de l'échantillon de l'ESCC sont toutefois survenus pour ces périodes de collecte, de sorte que les répondants de la base téléphonique de deux régions sociosanitaires (Muskoka-Parry Sound et Northwestern en Ontario) n'ont pas eu la chance d'être sélectionnés dans la partie A de l'EASS. Les poids de ces répondants qui n'ont pas eu la chance d'être choisis ont été redistribués aux autres répondants de l'ESCC provenant de la base aréolaire de ces deux régions sociosanitaires. Dans chacune des deux régions sociosanitaires, le facteur d'ajustement a été calculé comme suit :

$$\frac{\text{Somme des poids 3 pour tous les répondants de l' ESCC}}{\text{Somme des poids 3 pour tous les répondants de l' ESCC provenant de la base aréolaire}}$$

Pour les répondants de l'ESCC des autres régions sociosanitaires, le facteur d'ajustement a été fixé à 1. Le poids 3 a été multiplié par ce facteur d'ajustement pour produire le poids 3.1. Les répondants de l'ESCC provenant de la base téléphonique dans les deux régions problématiques ont été éliminés du processus de pondération à partir de ce point.

Ajustement 3.2 – Sous-échantillonnage de la partie A

À ce point, tous les répondants de l'ESCC étaient éligibles à être sélectionnés dans le sous-échantillon de la partie A de l'EASS. Ainsi, un sous-échantillon a été sélectionné et un ajustement au poids a été nécessaire pour refléter cette sélection. Étant donné que la sélection du sous-échantillon a été faite à l'échelle de la base de sondage dans chaque région sociosanitaire, le facteur d'ajustement a été calculé selon la base de sondage. Le facteur d'ajustement, qui représente l'inverse de la probabilité de sélection, a été calculé comme suit :

Somme des poids 3.1 pour tous les répondants de l' ESCC

Somme des poids 3.1 pour tous les personnes sélectionnées dans la partie A de l' EASS

Le poids 3.1 des personnes sélectionnées dans l'échantillon de la partie A de l'EASS a été multiplié par ce facteur d'ajustement pour produire le poids 3.2.

Ajustement 3.3 – Non-réponse

Même si les unités restantes sont tous des répondants de l'ESCC et que l'interview de l'EASS a été faite en même temps que l'interview de l'ESCC, certains d'entre eux ne peuvent être considérés répondants pour l'EASS. Pour être considéré répondant pour l'EASS, la personne devait avoir répondu à tous les modules de l'EASS. Par contre, cela n'a pas été le cas pour certains des répondants de l'ESCC tels que les répondants par procuration (les modules de l'EASS n'étaient pas demandés si l'interview était par procuration) et certains des répondants partiels à l'ESCC. Pour cette raison, les personnes sélectionnées dans l'EASS qui ne remplissent pas les critères ci-dessus ont été considérées comme non-répondants pour l'EASS. Par conséquent, le poids de ceux considérés non-répondants a été redistribué aux répondants de l'EASS à l'aide de classes de réponse.

À cette étape, les classes d'ajustement peuvent être définies à l'échelle de la province. Mais, étant donné que ce ne sont pas toutes les régions sociosanitaires qui sont couvertes dans la partie B de l'EASS (voir l'ajustement 4.1), et qu'il doit y avoir une intégration des poids des deux parties, les classes d'ajustement doivent être définies à l'intérieur de deux groupes dans chaque province. Le premier groupe contient toutes les régions sociosanitaires couvertes dans les deux parties de l'EASS. Le deuxième groupe contient les autres régions. Pour chacun des deux groupes et ce, à l'échelle de la province, des classes de réponse ont été créées. L'algorithme CHAID (Chi-Square Automatic Interaction Detector), disponible dans Knowledge Seeker², permet d'identifier les caractéristiques qui divisent le mieux l'échantillon en classes selon leurs propensions à répondre. En fonction du groupe et de la province, les caractéristiques suivantes ont été utilisées pour créer les classes d'ajustement : restriction des activités, difficultés dans les situations sociales, problème de l'ouïe, genre de fumeur, auto-évaluation de la santé mentale, sexe, âge, éducation, et la taille du ménage. Dans chaque classe, un ajustement a donc été calculé de la façon suivante :

Somme des poids 3.2 pour tous les personnes sélectionnées dans la partie A de l' EASS

Somme des poids 3.2 pour tous les personnes répondantes à la partie A de l' EASS

Le poids 3.2 des répondants à la partie A de l'EASS a été multiplié par ce facteur d'ajustement pour produire le poids 3.3. Les personnes non-répondantes ont été éliminées du processus de pondération à partir de ce point.

Ces poids sont les poids finaux de la partie A de l'EASS; ils seront intégrés aux poids finaux de la partie B de l'EASS.

² ANGOSS Software (1995). Knowledge Seeker IV for Windows - User's Guide. ANGOSS Software International Limited.

Ajustement 4 – Étapes d'ajustement pour la partie B

En ce qui concerne les ajustements s'appliquant seulement à la partie B de l'EASS, seuls les répondants de l'ESCC des périodes de collecte de janvier à août 2003 ont été conservés. Pour ceux-ci, le poids 2 est maintenant appelé le poids 4. Les ajustements 4.1 à 4.8 ci-dessous s'appliquent seulement à la partie B de l'EASS.

Ajustement 4.1 – Régions sociosanitaires non couvertes

Suite à la répartition de l'échantillon de la partie B de l'EASS, il y avait 21 régions sociosanitaires pour lesquelles aucun échantillon n'avait été alloué; ces régions n'étaient donc pas couvertes pour cette partie. Les répondants de l'ESCC des périodes de collecte de janvier à août 2003 dans ces régions ont donc été éliminés du processus à partir de ce point. Le poids 4 des répondants des autres régions est maintenant appelé le poids 4.1. À partir de cette étape, la partie B de l'EASS couvre seulement la population de 109 régions sociosanitaires (130 – 21).

Ajustement 4.2 – Bases de sondage non couvertes

Dans la partie A de l'EASS, il y avait également 31 autres régions sociosanitaires (des 109 régions couvertes) où aucun échantillon n'avait été alloué aux répondants de l'ESCC provenant de la base aréolaire. Alors, pour chacune de ces régions, le poids des répondants de l'ESCC provenant de la base aréolaire a été redistribué aux répondants de l'ESCC provenant de la base téléphonique. Dans chacune de ces 31 régions sociosanitaires, l'ajustement suivant a été calculé :

$$\frac{\text{Somme des poids 4.1 pour tous les répondants de l' ESCC}}{\text{Somme des poids 4.1 pour tous les répondants de l' ESCC provenant de la base téléphonique}}$$

Pour les répondants de l'ESCC des autres régions, l'ajustement a été fixé à 1. Le poids 4.1 des répondants dans les bases de sondage couvertes a été multiplié par ce facteur d'ajustement pour produire le poids 4.2. Les répondants de l'ESCC provenant de la base aréolaire des 31 régions mentionnées ont été éliminés du processus de pondération à partir de ce point.

Ajustement 4.3 – Exclusion des répondants du sous-échantillon 3 de l'ESCC

Le sous-échantillon 3 de l'ESCC représente un sous-ensemble des répondants de l'ESCC à partir de la période de collecte du mois d'avril (avant avril, il n'y avait pas de sous-échantillon 3). Les unités du sous-échantillon 3 ont reçu les modules « Satisfaction des patients » et « Accès aux services de soins de santé ». Puisque la partie B de l'EASS consiste principalement à demander le module « Accès aux services de soins de santé » et le module sur les temps d'attentes, il a alors été décidé d'exclure de la partie B de l'EASS, les répondants de l'ESCC qui ont déjà répondu au module « Accès aux services de soins de santé » (i.e., dans le sous-échantillon 3 de l'ESCC). À l'intérieur de chaque combinaison base de sondage/région sociosanitaire, les poids des répondants du sous-échantillon 3 de l'ESCC ont été distribués aux répondants qui n'étaient pas dans le sous-échantillon 3 de l'ESCC (entre avril et août 2003). Le facteur d'ajustement a été calculé comme suit :

$$\frac{\text{Somme des poids 4.2 pour tous les répondants de l' ESCC}}{\text{Somme des poids 4.2 pour tous les répondants de l' ESCC ne faisant pas partie du sous - échantillon 3}}$$

Pour les répondants de l'ESCC avant avril 2003, le facteur d'ajustement a été fixé à 1. Le nouveau poids 4.3 est égal au poids 4.2 multiplié par ce facteur d'ajustement. Les répondants du sous-échantillon 3 de l'ESCC ont été éliminés du processus de pondération à partir de ce point.

Ajustement 4.4 - Exclusion des répondants provenant des logements multiples

Cet ajustement s'applique seulement aux répondants de l'ESCC provenant de la base aréolaire.

Pour des raisons techniques, les répondants de l'ESCC qui proviennent de logements multiples n'ont eu aucune chance de faire partie de l'échantillon de l'EASS, ils ont été exclus du processus de sélection de l'échantillon de la partie B. À l'échelle de la région sociosanitaire, leurs poids ont donc été distribués aux autres répondants de la base aréolaire. Le facteur d'ajustement a été calculé comme suit :

$$\frac{\text{Somme des poids 4.3 pour tous les répondants de l'ESCC}}{\text{Somme des poids 4.3 pour tous les répondants de l'ESCC provenant de logements multiples}}$$

Le poids 4.3 des répondants de l'ESCC ne provenant pas de logements multiples a été multiplié par ce facteur d'ajustement pour produire le poids 4.4. Les répondants de l'ESCC provenant de logements multiples ont été éliminés du processus de pondération à partir de ce point.

Ajustement 4.5 – Exclusion des répondants par procuration

Puisque les modules de l'EASS ne peuvent être demandés par procuration, il était préférable de ne pas recontacter les répondants par procuration de l'ESCC pour la partie B de l'EASS. Ainsi, les poids des répondants par procuration ont été distribués aux autres répondants en utilisant des classes d'ajustement. Tout comme pour l'ajustement 3.3 (non-réponse à la partie A), l'algorithme CHAID a été utilisé pour identifier les caractéristiques qui divisent le mieux l'échantillon en classes selon leurs propensions à répondre par procuration ou non. Les classes ont été définies à l'échelle de la région sociosanitaire par base de sondage. Les caractéristiques suivantes ont été utilisées pour définir les classes : taille du ménage, restriction des activités, difficultés avec les situations sociales, langue(s) parlée(s), statut d'emploi, indicateur d'immigration, revenu, éducation, âge, sexe et quelques autres caractéristiques. Un facteur d'ajustement a été calculé de la façon suivante à l'intérieur de chaque classe :

$$\frac{\text{Somme des poids 4.4 pour tous les répondants de l'ESCC}}{\text{Somme des poids 4.4 pour tous les répondants de l'ESCC qui n'ont pas été interviewés par procuration}}$$

Le poids 4.4 des répondants de l'ESCC qui n'ont pas été interviewés par procuration a été multiplié par ce facteur d'ajustement pour produire le poids 4.5. Les répondants par procuration de l'ESCC ont été éliminés du processus de pondération à partir de ce point.

Ajustement 4.6 - Exclusion des répondants avec un numéro de téléphone invalide

Cet ajustement s'applique seulement aux répondants de l'ESCC provenant de la base aréolaire.

La partie B de l'EASS a été réalisée en recontactant les répondants de l'ESCC à l'aide des numéros de téléphone fournis lors de l'interview de l'ESCC. Les répondants qui ont refusé de fournir leurs numéros de téléphone, ou qui ont fourni un numéro de téléphone invalide, ont été

exclus du processus de sélection pour la partie B de l'EASS. Leurs poids ont été distribués aux répondants de l'ESCC qui ont fourni des numéros de téléphone valides. Une fois de plus, l'algorithme CHAID a été utilisé pour identifier des classes d'ajustement à l'échelle de la région sociosanitaire. Les caractéristiques suivantes ont principalement été utilisées pour former les classes : le revenu du ménage, le statut propriétaire/locataire et la taille du ménage. Pour chaque classe, le facteur d'ajustement a été calculé comme suit :

$$\frac{\text{Somme des poids 4.5 pour tous les répondants de l' ESCC}}{\text{Somme des poids 4.5 pour tous les répondants de l' ESCC qui ont fourni un \# de téléphone}}$$

Le poids 4.5 des répondants de l'ESCC provenant de la base aréolaire qui ont fourni un numéro de téléphone a été multiplié par ce facteur d'ajustement pour produire le poids 4.6. Les répondants de l'ESCC provenant de la base aréolaire qui n'ont pas fourni un numéro de téléphone valide ont été éliminés du processus de pondération à partir de ce point.

Ajustement 4.7 – Sous-échantillonnage de la partie B

À partir de ce moment-ci, tous les répondants de l'ESCC restants étaient éligibles à être sélectionnés pour la partie B de l'EASS. Pour chaque combinaison région sociosanitaire/base de sondage, un sous-échantillon des répondants de l'ESCC a été sélectionné. Un facteur d'ajustement reflétant les probabilités de sélection a été calculé comme suit :

$$\frac{\text{Somme des poids 4.6 pour tous les répondants de l' ESCC}}{\text{Somme des poids 4.6 pour toutes les personnes sélectionnées dans la partie B de l' EASS}}$$

Le poids 4.6 des personnes sélectionnées dans la partie B de l'EASS a été multiplié par ce facteur d'ajustement pour produire le poids 4.7. Les répondants de l'ESCC, qui n'ont pas été sélectionnés dans la partie B de l'EASS, ont été éliminés du processus à partir de ce point.

Ajustement 4.8 – Non-réponse

Lorsque l'intervieweur ne pouvait rejoindre la personne sélectionnée, ou que celle-ci refusait de participer à l'EASS, cette personne était considérée non-répondante à l'EASS et était rejetée pour l'enquête. À l'intérieur de chaque province, les poids de ces non-répondants à l'EASS ont été distribués aux répondants de l'EASS. L'algorithme CHAID a été utilisé pour identifier les caractéristiques qui divisent le mieux l'échantillon en classes selon leurs propensions à répondre. Les caractéristiques suivantes, obtenues lors de l'interview de l'ESCC, ont été principalement utilisées pour définir les classes : auto-évaluation de la santé générale, auto-évaluation de la santé mentale, restrictions des activités, incapacité au cours des deux dernières semaines, genre de fumeur, langue(s) parlée(s), état matrimonial, âge, sexe et le revenu. Pour chaque classe, le facteur d'ajustement suivant a été calculé :

$$\frac{\text{Somme des poids 4.7 pour toutes les personnes sélectionnées dans la partie B de l' EASS}}{\text{Somme des poids 4.7 pour toutes les personnes répondantes à la partie B de l' EASS}}$$

Le poids 4.7 des personnes répondantes à l'EASS a été multiplié par ce facteur d'ajustement pour produire le poids 4.8. Les personnes non-répondantes à l'EASS ont été éliminées du processus de pondération à partir de ce point.

Ces poids sont les poids finaux de la partie B de l'EASS; ils seront intégrés aux poids finaux de la partie A de l'EASS à la prochaine étape.

Ajustement 5 – Intégration des parties A et B

À ce point, deux séries de poids pour l'EASS ont été créées. Les poids finaux de la partie A représentent la population canadienne au complet (couverte par l'enquête), tandis que les poids finaux de la partie B représentent la population canadienne moins la population de 21 régions sociosanitaires. L'objectif de cet ajustement est d'intégrer les poids des répondants des deux parties afin de créer un seul poids de façon à ce que les poids intégrés représentent la population canadienne au complet. Voici de quelle façon l'intégration a été faite.

Premièrement, les poids des répondants de la partie A dans les 21 régions sociosanitaires non couvertes dans la partie B n'ont pas eu besoin d'être ajustés. Pour ceux-ci, le poids intégré (poids 5) est égal au poids 3.3. Seulement les poids des répondants des régions couvertes par les deux parties ont été ajustés. Pour ces répondants, un facteur d'ajustement entre 0 et 1 a été déterminé de façon à représenter l'importance relative de chaque échantillon dans l'échantillon total. Cette importance relative a été mesurée en termes de taille d'échantillon. Plus la proportion d'échantillon qu'une partie représente dans l'échantillon total est grande, plus grande sera son importance relative dans l'échantillon total. Pour obtenir le facteur d'ajustement d'intégration, dans chaque province, on calcule d'abord un facteur α obtenu de la façon suivante :

$$\alpha = \frac{n_A}{n_A + n_B}$$

où n_A et n_B représentent la taille d'échantillon de la partie A et de la partie B respectivement dans chaque province. Le poids des répondants de la partie A a été multiplié par ce facteur α , tandis que le poids des répondants de la partie B a été multiplié par $1 - \alpha$. Le produit du facteur d'ajustement dérivé ici, par le poids final calculé auparavant (3.3 ou 4.8 tout dépendant de quelle partie provient l'unité), procure le poids intégré 5.

Noter que suite à la série d'ajustements appliqués sur les poids, il est possible que certaines unités se retrouvent avec des poids se démarquant des autres poids de leur province au point même de devenir aberrants. Certains répondants peuvent effectivement représenter une proportion anormalement élevée de leur province et ainsi influencer fortement les estimations de ces provinces. Afin d'éviter cette situation, le poids des répondants qui contribuent de façon aberrante à leur province est « winsorisé » à la baisse.

Ajustement 6 – Poststratification

La dernière étape consiste en une poststratification. Elle est appliquée afin d'assurer que la somme des poids finaux corresponde aux estimations de population de 2003 pour chaque groupe province-âge-sexe et pour chaque région métropolitaine de recensement (RMR). Les comptes de population et les groupes de contrôle suivants ont été utilisés pour cette étape :

1. Pour les hommes et les femmes, quatre groupes d'âge (15-24, 25-44, 45-64, 65+) ont été définis à l'intérieur de chaque province. Les comptes de population pour ces groupes de contrôle, étaient la moyenne des projections du recensement pour les mois de janvier à décembre 2003.

2. Aussi, les RMR (25) et les régions non-RMR dans chaque province ont été utilisées comme groupes de contrôle. Les comptes de population utilisés pour ces groupes, étaient également les projections du recensement pour les mois de janvier à décembre 2003.

La méthode d'estimation de régression généralisée (GREG) a été utilisée pour modifier les poids intégrés (poids 5) de façon à ce que les poids poststratifiés (poids 6), pour chaque groupe de contrôle ci-dessus, concordent avec les comptes projetés de population.

Poids maître de l'EASS

Suite à l'application de tous ces ajustements, le poids maître de l'EASS a été créé et il se retrouve sur le fichier de données sous le nom WTSC_HSM.

Ajustement 7 - Ajustement de partage

Dans le but de partager les données de santé avec les Ministères provinciaux de la santé ainsi qu'avec Santé Canada, un fichier contenant seulement les répondants qui ont accepté de partager leurs données avec ces organismes a été créé. Les répondants qui ont refusé de partager leurs données ont été éliminés du processus à cette étape et leurs poids ont été distribués aux autres répondants qui ont accepté de partager leurs données en utilisant des classes d'ajustement. Les classes ont été définies à l'échelle de la province en utilisant l'algorithme CHAID. Dans toutes les provinces, la question concernant le revenu personnel a été utilisée comme première variable pour créer les classes. D'autres caractéristiques telles que l'éducation, l'âge, le sexe, l'indicateur d'immigration et l'auto-évaluation de la santé générale ont aussi été utilisées pour définir les classes. Dans chaque classe, le facteur d'ajustement a été calculé comme suit :

$$\frac{\text{Somme des poids 6 pour toutes les personnes répondantes à l' EASS}}{\text{Somme des poids 6 pour tous les répondants de l' EASS ayant accepté de partager}}$$

Le poids 6 des répondants qui ont accepté de partager a été multiplié par ce facteur d'ajustement pour produire le poids 7. Les répondants qui ont refusé de partager ont été enlevés du fichier à partir de ce point.

Ajustement 8 – Poststratification après l'ajustement de partage

Suite à l'ajustement de partage, la somme du poids 7 dans chaque groupe province-âge-sexe et RMR ne concordait plus exactement avec les comptes de population. Alors, une poststratification a été appliquée à nouveau afin que la somme des poids de partage finaux concorde avec les comptes de population projetés pour 2003. Tout comme à l'ajustement 6, la méthode d'estimation par régression généralisée (GREG) a été utilisée pour produire les poids de partage finaux (poids 8).

Poids de partage de l'EASS

Suite à l'application de tous ces ajustements, le poids de partage de l'EASS a été créé et il se retrouve sur le fichier de données sous le nom WTSC_HSS.

8. Qualité des données

8.1 Taux de réponse

Compte tenu du plan d'échantillonnage complexe de l'EASS, il n'est pas possible de rapporter un taux de réponse global pour l'EASS. On peut seulement rapporter le taux de réponse pour chacune des deux parties de l'échantillon (A et B) indépendamment.

Pour la partie A, après avoir retiré les unités hors du champ de l'enquête, 26 394 ménages ont été sélectionnés pour participer à l'ESCC/EASS. De ce nombre, 22 972 ont accepté de participer à l'enquête ce qui résulte en un taux de réponse à l'échelle du ménage de 87,0 %. Parmi ces ménages répondants, 21 680 personnes ont été sélectionnées pour participer à l'enquête et faisaient partie de la population cible de l'EASS (les 12 à 14 ans sont exclus de l'EASS). Parmi ces personnes sélectionnées, 18 981 ont été considérées répondantes pour l'EASS, ce qui résulte en un taux de réponse à l'échelle de la personne de 87,6 %. Une personne était considérée répondante pour l'EASS, si elle avait répondu aux modules du questionnaire de l'EASS (une personne pouvait être considérée répondante pour l'ESCC mais pas pour l'EASS). Ainsi, à l'échelle nationale, un taux de réponse combiné de **76,2 %** a donc été observé pour la partie A de l'EASS. Le taux de réponse combiné s'obtient en multipliant les taux de réponse aux échelles du ménage et de la personne. Les taux de réponse à l'échelle provinciale sont présentés dans le tableau 8.1.

Pour la partie B, 15 051 personnes avaient été sélectionnées (parmi les répondants de l'ESCC) pour participer à l'EASS. De ce nombre, 13 024 ont accepté de participer à l'enquête, ce qui résulte en un taux de réponse de **86,5 %** à l'échelle nationale pour la partie B. Il est à noter que toutes les personnes sélectionnées pour la partie B (15 051) ont été considérées dans le champ de l'enquête pour le calcul du taux de réponse même s'il était impossible de rejoindre la personne (par exemple, à cause d'un mauvais numéro de téléphone). Les taux de réponse à l'échelle provinciale sont présentés au tableau 8.1.

Nous décrivons, dans ce qui suit, de quelle façon les différentes composantes de l'équation doivent être manipulées afin de calculer correctement les taux de réponse combinés pour la partie A de l'EASS.

Taux de réponse à l'échelle du ménage

$$\text{HHRR} = \frac{\text{\# de ménages répondants}}{\text{tous les ménages faisant partie du champ de l'enquête}}$$

Taux de réponse à l'échelle de la personne

$$\text{PPRR} = \frac{\text{\# de répondants}}{\text{toutes les personnes sélectionnées faisant partie du champ de l'enquête}}$$

$$\text{Taux de réponse combiné (partie A)} = \text{HHRR} \times \text{PPRR}$$

Voici maintenant un exemple de calcul du taux de réponse combiné pour la partie A de l'EASS pour le Canada en utilisant l'information fournie au tableau 8.1.

$$\text{HHRR} = \frac{22\,972}{26\,394} = 0,870$$

$$\text{PPRR} = \frac{18\,981}{21\,680} = 0,876$$

$$\begin{aligned} \text{Taux de réponse combiné} &= 0,870 \times 0,876 \\ &= 0,762 \\ &= \mathbf{76,2\%} \end{aligned}$$

Tableau 8.1 Taux de réponse pour la partie A et B de l'EASS

Province	Partie A							Partie B		
	Ménage			Personne			Taux de rép. comb.	# pers. cibles	# pers. rép.	Taux de rép.
	# Mén. cibles	# Mén. rép.	Taux de rép. mén.	# pers. rép.	# pers. rép.	Taux de rép. pers.				
Terre Neuve-et-Labrador	1 065	991	93,1	934	826	88,4	82,3	1 934	1 694	87,6
Île-du Prince Édouard	744	653	87,8	622	559	89,9	78,9	801	691	86,3
Nouvelle Écosse	1 397	1 245	89,1	1 174	1 055	89,9	80,1	2 160	1 875	86,8
Nouveau Brunswick	1 327	1 224	92,2	1 149	1 025	89,2	82,2	2 281	1 870	82,0
Québec	4 038	3 373	83,5	3 204	2 828	88,3	73,7	1 216	1 048	86,2
Ontario	5 363	4 442	82,8	4 192	3 505	83,6	69,2	1 225	1 001	81,7
Manitoba	2 063	1 877	91,0	1 767	1 591	90,0	81,9	1 933	1 734	89,7
Saskatchewan	2 213	1 992	90,0	1 876	1 661	88,5	79,7	1 745	1 575	90,3
Alberta	3 776	3 357	88,9	3 146	2 700	85,8	76,3	776	652	84,0
Colombie Britannique	4 408	3 818	86,6	3 616	3 231	89,4	77,4	980	884	90,2
Canada	26 394	22 972	87,0	21 680	18 981	87,6	76,2	15 051	13 024	86,5

Le tableau 8.2 présente le nombre total de répondants obtenus pour l'EASS (partie A et partie B).

Tableau 8.2 Nombre total de répondants pour l'EASS par province

Province	Nombre de répondants
Terre-Neuve-et-Labrador	2 520
Île-du-Prince-Édouard	1 250
Nouvelle-Écosse	2 930
Nouveau-Brunswick	2 895
Québec	3 876
Ontario	4 506
Manitoba	3 325
Saskatchewan	3 236
Alberta	3 352
Colombie-Britannique	4 115
Canada	32 005

8.2 Erreurs dans les enquêtes

L'enquête permet de produire des estimations fondées sur l'information recueillie à partir d'un échantillon de personnes. On aurait pu obtenir des estimations quelque peu différentes si on avait effectué un recensement complet en utilisant le même questionnaire, les mêmes intervieweurs, les mêmes superviseurs, les mêmes méthodes de traitement, etc., que ceux utilisés pour l'enquête. La différence entre les estimations tirées de l'échantillon et celles qui découlent d'un dénombrement complet effectué dans des conditions semblables s'appelle l'erreur due à l'échantillonnage des estimations.

Les erreurs qui ne sont pas liées à l'échantillonnage peuvent être commises à presque toutes les étapes d'une enquête. Il est possible que les intervieweurs comprennent mal les instructions, que les répondants fassent des erreurs en complétant le questionnaire, que les réponses soient mal saisies et que des erreurs se produisent au moment du traitement et de la totalisation des données. Tous ces exemples représentent des erreurs non dues à l'échantillonnage.

8.2.1 Erreurs non dues à l'échantillonnage

Sur un grand nombre d'observations, les erreurs aléatoires auront peu d'effet sur les estimations tirées de l'enquête. Toutefois, les erreurs qui se produisent systématiquement contribueront à des biais dans les estimations de l'enquête. On a consacré beaucoup de temps et d'efforts à réduire les erreurs non dues à l'échantillonnage dans l'enquête. Des mesures d'assurance de la qualité ont été appliquées à chaque étape du cycle de collecte et de traitement des données afin de contrôler la qualité des données. On a notamment fait appel à des intervieweurs hautement qualifiés, une formation poussée sur les méthodes d'enquête et le questionnaire et l'observation des intervieweurs afin de déceler les problèmes. La mise à l'essai de l'application IAO et les essais sur le terrain ont également été au nombre des procédures essentielles pour réduire au maximum les erreurs de collecte de données.

L'effet de la non-réponse sur les résultats de l'enquête constitue une source importante d'erreurs non dues à l'échantillonnage dans les enquêtes. L'ampleur de la non-réponse varie de non-réponse partielle (le fait de ne pas répondre à une ou plusieurs questions) à une non-réponse totale. Dans le cas de l'EASS, il n'y a presque pas eu de non-réponse partielle car une fois le questionnaire débuté les répondants avaient tendance à le terminer. Il y a eu non-réponse totale lorsque la personne sélectionnée pour participer à l'enquête a refusé de le faire ou que l'intervieweur a été incapable d'entrer en contact avec elle. On a traité les cas de non-réponse totale en corrigeant les poids des personnes qui ont répondu à l'enquête afin de compenser pour celles qui n'ont pas répondu. Voir la section 8 pour avoir de plus amples détails sur la correction de la pondération pour la non-réponse.

8.2.2 Erreurs dues à l'échantillonnage

Étant donné que les estimations d'une enquête par sondage comportent inévitablement des erreurs dues à l'échantillonnage, de bonnes méthodes statistiques exigent que les chercheurs fournissent aux utilisateurs une certaine indication de l'ampleur de cette erreur. La mesure de l'importance éventuelle des erreurs dues à l'échantillonnage est fondée sur l'écart type des estimations tirées des résultats de l'enquête. Cependant, en raison de la grande diversité des estimations que l'on peut tirer d'une enquête, l'écart type d'une estimation est habituellement exprimé en fonction de l'estimation à laquelle il se rapporte. La mesure résultante, appelée coefficient de variation (CV), s'obtient en divisant l'écart type de l'estimation par l'estimation elle-même et on l'exprime en pourcentage de l'estimation.

Par exemple, supposons qu'une personne estime que 25 % des Canadiens âgés de 15 ans et plus ont eu besoin de consulter un médecin spécialiste dans la dernière année et que cette estimation comporte un écart type de 0,012. On calcule alors le CV de cette estimation de la façon suivante :

$$(0,012/0,25) \times 100 \% = 4,80 \%$$

Statistique Canada utilise fréquemment les résultats du CV pour l'analyse des données et conseille vivement aux utilisateurs produisant des estimations à partir des fichiers de données de l'EASS de faire de même.

8.2.2.1 Méthode bootstrap pour l'estimation de la variance

Afin de déterminer la qualité de l'estimation et de calculer le CV, il est nécessaire de calculer l'écart type. Les intervalles de confiance nécessitent également le calcul de l'écart type de l'estimation.

Pour l'EASS, on utilise un plan de sondage à plusieurs degrés, ce qui signifie qu'il n'y a pas de formule simple pour calculer les estimations de la variance. Par conséquent, il a fallu appliquer une méthode approximative. On applique la méthode bootstrap parce que l'information sur le plan d'échantillonnage doit être prise en compte lors du calcul des estimations de la variance. La méthode bootstrap permet d'effectuer ce calcul et, conjointement avec le programme Bootvar décrit à la sous-section suivante, elle demeure une méthode relativement facile à utiliser.

La méthode de ré-échantillonnage bootstrap utilisée pour l'EASS sous-tend la sélection d'échantillons aléatoires simples, appelés les échantillons répétés, et le calcul de l'écart des

estimations d'un échantillon répété à l'autre. Dans chaque strate, on sélectionne un échantillon aléatoire simple de $(n-1)$ des n grappes avec remise pour former un échantillon répété. Il est à noter que puisque la sélection se fait avec remplacement, une grappe peut être choisie plus d'une fois. Pour chaque échantillon répété, on recalcule le poids de sondage de chaque enregistrement dans les grappes $(n-1)$ sélectionnées. Ces poids sont par la suite poststratifiés en fonction de l'information démographique de la même façon que la pondération du plan de sondage, ce qui permet d'obtenir les poids bootstrap finaux.

Le processus complet (sélectionner les échantillons aléatoires simples, recalculer et poststratifier les poids de chaque strate) est répété B fois, B prenant une grande valeur. En général, on utilise $B=500$ pour l'EASS afin de produire 500 poids bootstrap. Pour obtenir l'estimateur bootstrap de la variance, on doit calculer l'estimation ponctuelle de chacun des B échantillons. L'écart type de ces estimations représente l'estimateur bootstrap de la variance. Statistique Canada a élaboré un programme qui peut effectuer tous ces calculs pour l'utilisateur : le programme Bootvar.

8.2.2.2 Programme Bootvar pour l'estimation de la variance

Le programme Bootvar est offert en format SAS et SPSS. Il est constitué de macros qui calculent les variances de totaux, ratios, différences entre ratios, et pour des régressions linéaires et logistiques. Le programme Bootvar est fourni avec les poids bootstrap et un document qui explique comment modifier et utiliser le programme selon les besoins de l'utilisateur.

9. Lignes directrices pour la totalisation, l'analyse et la diffusion

Cette section du guide décrit les lignes directrices que doivent suivre les utilisateurs qui totalisent, analysent, publient ou diffusent de quelque façon des données provenant des fichiers de microdonnées de l'enquête. Ces lignes directrices devraient leur permettre de reproduire les chiffres déjà publiés par Statistique Canada et de produire aussi des chiffres non encore publiés conformes aux lignes directrices établies.

9.1 Lignes directrices pour l'arrondissement

Afin que les estimations calculées d'après ces fichiers de microdonnées en vue d'être publiées ou diffusées de toute autre façon correspondent à celles produites par Statistique Canada, il est vivement conseillé à l'utilisateur de les arrondir en se conformant aux lignes directrices suivantes.

- a) Les estimations qui figurent dans le corps d'un tableau statistique doivent être arrondies à la centaine près par la méthode d'arrondissement classique. Selon cette méthode, si le premier ou le seul chiffre à supprimer se situe entre 0 et 4, le dernier chiffre retenu ne change pas. Si le premier ou le seul chiffre à supprimer se situe entre 5 et 9, on augmente d'une unité (1) la valeur du dernier chiffre retenu. Par exemple, si l'on veut arrondir à la centaine près de la façon classique une estimation dont les deux derniers chiffres sont compris entre 00 et 49, il faut les remplacer par 00 et ne pas modifier le chiffre précédent (le chiffre des centaines). Si les deux derniers chiffres sont compris entre 50 et 99, il faut les remplacer par 00 et augmenter d'une unité (1) le chiffre précédent.
- b) Les totaux partiels de marge et les totaux de marge des tableaux statistiques doivent être calculés à partir de leurs éléments correspondants non arrondis, puis arrondis à leur tour à la centaine près selon la méthode d'arrondissement classique.
- c) Les moyennes, les proportions, les taux et les pourcentages doivent être calculés à partir d'éléments non arrondis (c'est-à-dire les numérateurs et (ou) dénominateurs), puis arrondis à une décimale par la méthode d'arrondissement classique. Si l'on veut arrondir une estimation à un seul chiffre décimal par cette méthode et que le dernier ou le seul chiffre à supprimer se situe entre 0 et 4, le dernier chiffre à retenir ne change pas. Si le premier ou le seul chiffre à supprimer se situe entre 5 et 9, on augmente d'une unité (1) le dernier chiffre à retenir.
- d) Les sommes et les différences d'agrégats (ou de rapports) doivent être calculées à partir de leurs éléments correspondants non arrondis, puis arrondies à leur tour à la centaine près (ou à la décimale près) selon la méthode d'arrondissement classique.
- e) Si, en raison de contraintes d'ordre technique ou autre, on applique une autre méthode que l'arrondissement classique, si bien que les estimations qui seront publiées ou diffusées de toute autre façon diffèrent des estimations correspondantes publiées par Statistique Canada, il est vivement conseillé à l'utilisateur d'indiquer la raison de ces divergences dans le ou les documents à publier ou à diffuser.
- f) Des estimations non arrondies ne doivent être publiées ou diffusées de toute autre façon en aucune circonstance. Des estimations non arrondies donnent l'impression d'être beaucoup plus précises qu'elles ne le sont en réalité.

9.2 Lignes directrices pour la pondération de l'échantillon en vue de la totalisation

Le plan d'échantillonnage utilisé pour l'EASS n'est pas autopondéré. Autrement dit, le poids d'échantillonnage n'est pas le même pour toutes les personnes qui font partie de l'échantillon. Même pour produire des estimations simples, y compris des tableaux statistiques ordinaires, l'utilisateur doit employer le poids d'échantillonnage approprié. Sinon, les estimations calculées à partir des fichiers de microdonnées ne pourront être considérées comme représentatives de la population observée et ne correspondront pas à celles de Statistique Canada. L'utilisateur ne doit pas non plus perdre de vue qu'en raison du traitement réservé à la zone du poids, certains logiciels ne permettent pas d'obtenir des estimations qui coïncident exactement avec celles de Statistique Canada.

9.2.1 Définitions des catégories d'estimations : de type nominal par opposition à quantitatives

Avant d'exposer la façon de totaliser et d'analyser les données du cycle 1.1 de l'EASS, il est bon de décrire les deux grandes catégories d'estimations ponctuelles des caractéristiques de la population qui peuvent être produites d'après le fichier de microdonnées de l'enquête.

Estimations de type nominal :

Les estimations de type nominal sont des estimations du nombre ou du pourcentage de personnes qui, dans la population visée par l'enquête, possèdent certaines caractéristiques ou font partie d'une catégorie particulière. Le nombre de personnes qui fument tous les jours est un exemple d'estimation de ce genre. L'estimation du nombre de personnes qui possèdent une caractéristique particulière peut aussi être appelée « estimation d'un agrégat ».

Exemple de question de type nominal :

WTMC_03

Avez-vous été référé par :

... votre médecin de famille?

... un autre spécialiste?

... un autre professionnel de la santé?

N'ai pas eu besoin d'être référé(e).

Estimations quantitatives :

Les estimations quantitatives sont des estimations de totaux ou de moyennes, de médianes ou d'autres mesures de tendance centrale de quantités qui ont trait à tous les membres de la population observée ou à certains d'entre eux.

Un exemple d'estimation quantitative est le nombre moyen de cigarettes que fument par jour les personnes qui fument tous les jours. Le numérateur correspond à l'estimation du nombre total de cigarettes que fument par jour les personnes qui fument tous les jours et le dénominateur, à l'estimation du nombre de personnes qui fument tous les jours.

Exemple de question quantitative :

WTM_Q07A Combien de temps avez-vous attendu entre le moment où [vous et votre médecin avez décidé que vous devriez rencontrer un spécialiste/vous et votre professionnel de la santé avez décidé que vous devriez rencontrer un spécialiste/le rendez-vous a été fait] et le moment où vous avez visité le spécialiste?

|_|_|_| (nombre d'unités)

9.2.2 Totalisation d'estimations de type nominal

On peut obtenir, à partir des fichiers de microdonnées, des estimations du nombre de personnes qui possèdent une caractéristique donnée en additionnant les poids finaux de tous les enregistrements contenant des données sur la caractéristique étudiée.

Pour obtenir les proportions et les rapports de la forme \hat{X} / \hat{Y} , on doit :

- a) additionner les poids finaux des enregistrements contenant la caractéristique voulue pour le numérateur (\hat{X});
- b) additionner les poids finaux des enregistrements contenant la caractéristique voulue pour le dénominateur (\hat{Y});
- c) diviser l'estimation du numérateur par celle du dénominateur.

9.2.3 Totalisation d'estimations quantitatives

Pour obtenir des estimations quantitatives d'après le fichier de microdonnées, on doit :

- a) multiplier la valeur de la variable étudiée par le poids final, puis faire la somme de cette quantité pour tous les enregistrements visés afin d'obtenir le numérateur (\hat{X});
- b) faire la somme des poids finaux des enregistrements contenant la variable étudiée pour obtenir le dénominateur (\hat{Y});
- c) diviser l'estimation du numérateur par l'estimation du dénominateur.

9.3 Lignes directrices pour l'analyse statistique

L'EASS se fonde sur un plan de sondage complexe qui prévoit une stratification et un échantillonnage à plusieurs degrés, ainsi que la sélection des répondants avec probabilités inégales. L'utilisation des données provenant d'une enquête aussi complexe pose des difficultés aux analystes, car le choix des méthodes d'estimation et de calcul de la variance dépend du plan de sondage et des probabilités de sélection.

Nombre de méthodes d'analyse intégrées aux progiciels statistiques permettent d'utiliser des poids, mais la signification et la définition de ces poids peuvent différer de celles applicables dans le contexte d'une enquête par sondage. Par conséquent, si les estimations calculées au

moyen de ces progiciels sont souvent exactes, les variances n'ont, quant à elles, pratiquement aucune signification.

Dans le cas de nombreuses méthodes d'analyse (par exemple la régression linéaire, la régression logistique, l'analyse de la variance), on peut rendre l'application des progiciels courants plus significative en rééchelonnant les poids qui figurent dans les enregistrements de façon à ce que le poids moyen soit égal à un (1). Les résultats produits par les progiciels classiques sont ainsi plus raisonnables puisque, même s'ils ne reflètent toujours pas la stratification et la mise en grappes du plan d'échantillonnage, ils tiennent compte de la sélection avec probabilités inégales. On peut effectuer cette transformation en utilisant dans l'analyse un poids égal au poids original divisé par la moyenne des poids originaux pour les unités échantillonnées (personnes) qui contribuent à l'estimation en question.

9.4 Lignes directrices pour la diffusion

Avant de diffuser et (ou) de publier des estimations tirées des fichiers, l'utilisateur doit déterminer le nombre de personnes qui ont fourni les données entrant dans le calcul de l'estimation. Si ce nombre est inférieur à 10, l'estimation pondérée ne doit pas être diffusée, quelle que soit la valeur de son coefficient de variation. Pour les estimations pondérées basées sur des échantillons d'au moins 10 personnes, l'utilisateur doit calculer le coefficient de variation de l'estimation arrondie et suivre les lignes directrices qui suivent.

Table 9.1 Lignes directrices relatives à la variabilité d'échantillonnage

Type d'estimation	CV (en %)	Lignes directrices
Acceptable	$0.0 \leq CV \leq 16.6$	On peut envisager une diffusion générale non restreinte des estimations. Aucune annotation particulière n'est nécessaire.
Marginale	$16.6 < CV \leq 33.3$	On peut envisager une diffusion générale non restreinte des estimations, en y joignant une mise en garde aux utilisateurs quant à la variabilité d'échantillonnage élevée liée aux estimations. Les estimations de ce genre doivent être identifiées par la lettre E (ou d'une autre manière similaire).
Inacceptable	$CV > 33.3$	<p>Statistique Canada recommande de ne pas publier des estimations dont la qualité est inacceptable. Toutefois, si l'utilisateur choisit de le faire, il doit alors adjoindre la lettre F (ou un autre identificateur semblable) et les diffuser avec l'avertissement suivant :</p> <p>« Nous avisons l'utilisateur que ...(précisez les données)... ne répondent pas aux normes de qualité de Statistique Canada pour ce programme statistique. Les conclusions tirées de ces données ne sauraient être fiables et seront fort probablement erronées. Ces données et toute conclusion qu'on pourrait en tirer ne doivent pas être publiées. Si l'utilisateur choisit de les publier, il est alors tenu de publier également le présent avertissement. »</p>

10. Utilisation du fichier

10.1 Utilisation des facteurs de pondération

Un seul poids WTMC_HSM figure dans le fichier de micro données à grande diffusion. Ce poids est applicable à chaque groupe d'âge, région, province ou territoire. TOUTES LES VARIABLES DU FICHER DEVRAIENT ÊTRE ANALYSÉES À L'AIDE DE CE POIDS. Pour une description plus détaillée du calcul de ce poids, consulter le section 8 sur la pondération.

10.2 Convention appliquée pour nommer les variables

On a adopté pour nommer les variables du cycle 1.1 de l'ESCC une convention qui permet aux utilisateurs des données de repérer et d'utiliser facilement celles-ci en fonction du module et du cycle. L'EASS a la même convention pour nommer les variables que l'ESCC. Les exigences qui suivent doivent être satisfaites : limiter les noms des variables à huit caractères au plus pour qu'il soit facile de les utiliser avec les logiciels d'analyse, préciser l'édition de l'enquête (cycle 1.1, 1.2...) dans le nom, et permettre de repérer facilement les variables conceptuellement identiques d'un cycle à l'autre de l'enquête. Les noms des variables correspondant à des modules ou à des questions identiques ne devraient différer qu'en ce qui concerne la position réservée dans le nom à l'identification du cycle particulier durant lequel les données ont été recueillies.

10.2.1 Structure élémentaire des noms des variables pour l'EASS

Chacun des huit caractères du nom d'une variable fournit des renseignements sur le type de données que contient la variable.

Positions 1 à 3 : Nom du module/de la section du questionnaire
Position 4 : Cycle de l'enquête
Position 5 : Type de variable
Positions 6 à 8 : Numéro de la question

Par exemple, la structure du nom de la variable correspondant à la question 6 de l'EASS 2003, module du temps d'attente, c'est-à-dire WTMC_06A, est la suivante :

Positions 1 à 3 : **WTM** Module du temps d'attente
Position 4 : **C** Cycle 2.1
Position 5 : **_** (_ = données recueillies)
Position 6 à 8 : **06A** numéro de la question et option de réponse

10.2.2 Positions 1 à 3 : Nom de la variable/section du questionnaire

On se sert des valeurs suivantes pour la composante du nom de la variable correspondant à la section du questionnaire :

ACC	Accès aux services de soins de santé	INC	Revenu
ADM	Admistration	LBF	Population active
CCC	Problèmes de santé chroniques	RAC	Limitation des activités
DHH	Variables démographiques et du ménage	SAM	Identificateurs d'échantillon
EDU	Éducation	SDC	Caractéristiques socio-démographiques
GEN	État de santé général	WTM	Temps d'attente
GEO	Identificateurs géographiques	WTS	Poids

10.2.3 Position 4 : Cycle

Cycle	Description
A	<p><u>Cycle 1.1 : Enquête sur la santé dans les collectivités canadiennes</u></p> <ul style="list-style-type: none"> ▪ enquête à l'échelle régionale, stratifiée selon la région socio-sanitaire ▪ contenu commun et contenu optionnel sélectionnés par les régions socio-sanitaires ▪ estimations à l'échelle des régions socio-sanitaires, des provinces, territoires et du Canada
B	<p><u>Cycle 1.2 : Enquête sur la santé dans les collectivités canadiennes, santé mentale et bien-être</u></p> <ul style="list-style-type: none"> ▪ enquête à l'échelle provinciale ▪ contenu thématique et contenu général supplémentaire ▪ estimations aux niveaux provincial, territorial et national
C	<p><u>Cycle 2.1 : Enquête sur la santé dans les collectivités canadiennes / Enquête sur l'Accès aux services de santé</u></p> <ul style="list-style-type: none"> ▪ enquête à l'échelle régionale, stratifiée selon la région socio-sanitaire ▪ contenu commun et contenu optionnel sélectionnés par les régions socio-sanitaires ▪ estimations à l'échelle des régions socio-sanitaires, des provinces, territoires et du Canada

10.2.4 Position 5 : Type de variable

_	Variable collectée	Variable qui figure directement sur le questionnaire.
C	Variable codée	Variable codée à partir d'une ou de plusieurs variables collectées (par exemple, code de la Classification type des industries (CTI)).
D	Variable dérivée	Variable calculée d'après une ou plusieurs variables collectés ou codées, ordinairement pendant le traitement au Bureau central (p. ex., indice de l'état de santé).
F	Variable indicatrice	Variable calculée à partir d'une ou de plusieurs variables collectées (comme variable dérivée), mais ordinairement par l'application informatique de collecte des données, aux fins de son utilisation ultérieure durant l'interview (p. ex., indicateur de travail).
G	Variable groupée	Variables collectées, codées, supprimées ou dérivées agrégées en un groupe (p. ex., groupe d'âge).

10.2.5 Positions 6 à 8 : Nom de la variable

En général, les trois dernières positions correspondent à la numérotation de la variable qui figure sur le questionnaire. On supprime la lettre « Q » utilisée pour représenter le mot "question" et on présente tous les numéros de question au moyen d'un groupe de deux chiffres. Par exemple, la question Q01A du questionnaire devient simplement 01A et la question 15 (Q15), simplement 15.

Parfois, certaines questions comportent plusieurs réponses alors la position finale dans la séquence du nom de la variable est représentée par une lettre. Pour ce genre de questions, de nouvelles variables sont créées dans le but de différencier un "oui" d'un "non" pour chaque possibilité de réponse. Par exemple, si la question Q2 a 4 réponses possibles, les nouvelles questions seraient Q2A pour la première possibilité, Q2B pour la deuxième, Q2C pour la troisième et ainsi de suite. Si seulement les options 2 et 3 sont choisies, alors Q2A = Non, Q2B = Oui, Q2C = Oui et Q2D = Non.

10.3 Accès aux fichiers maîtres

Une première approche offerte à tous les clients consiste à demander au personnel spécialisé des Services personnalisés à la clientèle de la Division de la statistique de la santé de produire des totalisations personnalisées. Ce service permet aux utilisateurs qui ne savent pas se servir de logiciels de totalisation d'obtenir des résultats personnalisés. Comme dans le cas du télé-accès, les résultats sont filtrés pour s'assurer qu'ils sont conformes aux normes de confidentialité et de fiabilité avant d'être diffusés. Contrairement au télé-accès, ce service est offert contre remboursement des frais. Vous pouvez contacter ce service en écrivant à : hd-ds@statcan.ca

Deuxièmement, le Programme des centres de données de recherche permet aux chercheurs de soumettre à Statistique Canada un projet de recherche fondé sur les données des fichiers

maîtres. Un ensemble particulier de règles est appliqué afin de décider quels projets seront acceptés. Lorsque le projet est accepté, le chercheur est considéré comme étant "réputé employé" par Statistique Canada pour la durée de l'étude et se voit accorder l'accès au fichier maître de l'enquête dans des locaux désignés de Statistique Canada. Veuillez consulter la page web suivante pour plus d'information :

http://www.statcan.ca/francais/rdc/index_f.htm