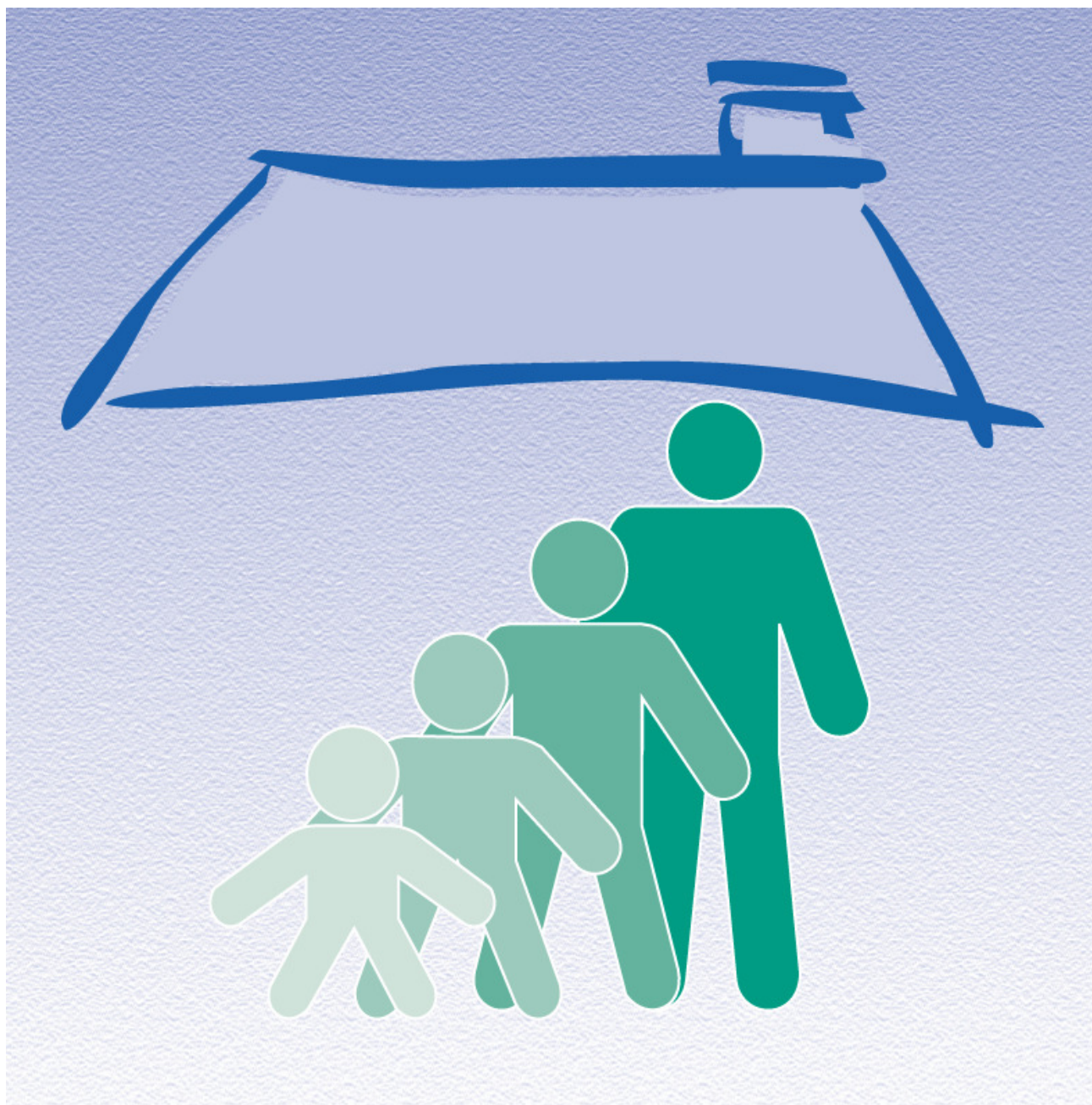


Enquête sur la santé dans les collectivités canadiennes (ESCC) – Composante annuelle

Guide de l'utilisateur Fichiers de microdonnées 2009

Jun 2010



Statistics
Canada

Statistique
Canada

Canada

Table des matières

1.	INTRODUCTION	1
2.	HISTORIQUE.....	3
3.	REMANIEMENT DE L'ESCC EN 2007	4
4.	CONTENU DE L'ESCC.....	6
4.1	CONTENU COMMUN	6
4.2	CONTENU OPTIONNEL	7
4.3	CONTENU DE RÉPONSE RAPIDE	7
4.4	CONTENU DES FICHIERS DE MICRODONNÉES	7
5.	PLAN D'ÉCHANTILLONNAGE	8
5.1	POPULATION CIBLE	8
5.2	RÉGIONS SOCIO-SANITAIRES	9
5.3	TAILLE ET RÉPARTITION DE L'ÉCHANTILLON	9
5.4	BASES DE SONDAGE ET STRATÉGIES D'ÉCHANTILLONNAGE DES MÉNAGES	11
5.4.1	ÉCHANTILLONNAGE DES MÉNAGES À PARTIR DE LA BASE ARÉOLAIRE.....	11
5.4.2	ÉCHANTILLONNAGE DES MÉNAGES À PARTIR DE LA BASE LISTE DE NUMÉROS DE TÉLÉPHONE.....	13
5.4.3	ÉCHANTILLONNAGE DES MÉNAGES À PARTIR DE LA BASE DE SONDAGE À CA DE NUMÉROS DE TÉLÉPHONE	14
5.5	RÉPARTITION DE L'ÉCHANTILLON PAR PÉRIODE DE COLLECTE DES DONNÉES	14
5.6	ÉCHANTILLONNAGE DES PERSONNES INTERVIEWÉES	15
5.7	ACHAT D'ÉCHANTILLON POUR L'ONTARIO	15
5.8	SOUS-ÉCHANTILLONNAGE POUR L'ENQUÊTE SUR L'ACCÈS AUX SERVICES DE SANTÉ (EASS)	16
6.	COLLECTE DES DONNÉES.....	18
6.1	INTERVIEWS ASSISTÉES PAR ORDINATEUR	18
6.2	DÉVELOPPEMENT DES APPLICATIONS DE L'ESCC	18
6.3	FORMATION DES INTERVIEWEURS.....	20
6.4	L'INTERVIEW	20
6.5	OPÉRATIONS SUR LE TERRAIN.....	23
6.6	CONTRÔLE DE LA QUALITÉ ET GESTION DE LA COLLECTE	23
7.	TRAITEMENT DES DONNÉES	25
7.1	VÉRIFICATION.....	25
7.2	CODAGE.....	25
7.3	CRÉATION DE VARIABLES DÉRIVÉES	25
7.4	PONDÉRATION	26
8.	PONDÉRATION.....	27
8.1	INTRODUCTION	27
8.2	PONDÉRATION DE L'ÉCHANTILLON PROVENANT DE LA BASE ARÉOLAIRE	28
8.3	PONDÉRATION DE L'ÉCHANTILLON PROVENANT DE LA BASE TÉLÉPHONIQUE	30
8.4	INTÉGRATION DES BASES ARÉOLAIRE ET TÉLÉPHONIQUE	32
8.5	LES ÉTAPES DE PONDÉRATION POST-INTÉGRATION	32
8.6	PARTICULARITÉS DE LA PONDÉRATION POUR LES TROIS TERRITOIRES	33
8.7	LA CRÉATION D'UN POIDS POUR LE FICHIER PARTAGÉ	34
9.	QUALITÉ DES DONNÉES	36
9.1	TAUX DE RÉPONSE	36
9.2	ERREURS DANS LES ENQUÊTES	37
9.2.1	ERREURS NON DUES À L'ÉCHANTILLONNAGE	37
9.2.2	ERREURS DUES À L'ÉCHANTILLONNAGE	38

10.	LIGNES DIRECTRICES POUR LA TOTALISATION, L'ANALYSE ET LA DIFFUSION	39
10.1	LIGNES DIRECTRICES POUR L'ARRONDISSEMENT	39
10.2	LIGNES DIRECTRICES POUR LA PONDÉRATION DE L'ÉCHANTILLON EN VUE DE LA TOTALISATION	40
10.2.1	DÉFINITIONS DES CATÉGORIES D'ESTIMATIONS : DE TYPE NOMINAL PAR OPPOSITION À QUANTITATIVES	40
10.2.2	TOTALISATION D'ESTIMATIONS DE TYPE NOMINAL.....	41
10.2.3	TOTALISATION D'ESTIMATIONS QUANTITATIVES	41
10.3	LIGNES DIRECTRICES POUR L'ANALYSE STATISTIQUE	42
10.4	LIGNES DIRECTRICES POUR LA DIFFUSION.....	42
11.	TABLEAUX DE LA VARIABILITÉ D'ÉCHANTILLONNAGE APPROXIMATIVE.....	44
11.1	COMMENT UTILISER LES TABLEAUX DE CV POUR LES ESTIMATIONS DE TYPE NOMINAL.....	44
11.2	EXEMPLES D'UTILISATION DES TABLEAUX DE CV POUR DES ESTIMATIONS DE TYPE NOMINAL.....	47
11.3	COMMENT UTILISER LES TABLEAUX DE CV POUR CALCULER LES LIMITES DE CONFIANCE	50
11.4	EXEMPLE D'UTILISATION DE TABLEAUX DE CV POUR OBTENIR DES LIMITES DE CONFIANCE.....	51
11.5	COMMENT UTILISER LES TABLEAUX DE CV POUR EFFECTUER UN TEST Z	51
11.6	EXEMPLE D'UTILISATION DES TABLEAUX DE CV POUR EFFECTUER UN TEST Z.....	52
11.7	VARIANCES OU COEFFICIENTS DE VARIATION EXACTS	52
11.8	SEUILS POUR LA DIFFUSION DES ESTIMATIONS RELATIVES À L'ESCC.....	53
12	FICHIERS DE MICRODONNÉES : DESCRIPTION, ACCÈS ET UTILISATION	54
12.1	FICHIERS MAÎTRES	54
12.1.1	CENTRE DE DONNÉES DE RECHERCHE	54
12.1.2	TOTALISATIONS PERSONNALISÉES	54
12.1.3.	TÉLÉ-ACCÈS.....	54
12.2	FICHIERS DE PARTAGE	55
12.3	FICHIERS DE MICRODONNÉES À GRANDE DIFFUSION	55
12.4	UTILISATION DES FICHIERS DE L'ESCC : FICHIER ANNUEL OU FICHIER DEUX ANS ?	57
12.5	UTILISATION DE LA VARIABLE DE PONDÉRATION.....	57
12.6	CONVENTION APPLIQUÉE POUR NOMMER LES VARIABLES À PARTIR DE 2007.....	58
12.7	CONVENTION APPLIQUÉE POUR NOMMER LES VARIABLES AVANT 2007	61
12.8	LIGNES DIRECTRICES POUR L'UTILISATION DES VARIABLES D'UN SOUS-ÉCHANTILLON – NE S'APPLIQUE PAS AUX FICHIERS DE 2009	61
12.8.2	ANALYSE COMBINANT LE CONTENU DES SOUS-ÉCHANTILLONS ET LE CONTENU OPTIONNEL	61
12.9	DICTIONNAIRES DE DONNÉES.....	61
12.10	DIFFÉRENCES DANS LE CALCUL DES VARIABLES DE CONTENU COMMUN FONDÉ SUR DIFFÉRENTS FICHIERS	62
	ANNEXE A	63
	ANNEXE A- ENQUÊTE SUR LA SANTÉ DANS LES COLLECTIVITÉS CANADIENNES (2009-2010)	65
	ANNEXE B.....	67
	ANNEXE B – SÉLECTION DU CONTENU OPTIONNEL PAR LES PROVINCES ET TERRITOIRES	69
	ANNEXE C	73
	ANNEXE C - GÉOGRAPHIE DISPONIBLE SUR LE FICHIER MAÎTRE ET DE PARTAGE ET CODES CORRESPONDANTS : CANADA, PROVINCES/TERRITOIRES, RÉGIONS SOCIO-SANITAIRES ET GROUPES HOMOLOGUES	75
	ANNEXE D	79
	ANNEXE D - RÉPARTITION DE L'ÉCHANTILLON PAR RÉGION SOCIO-SANITAIRE ET PAR BASE DE SONDAGE DE L'ESCC 2009	80
	ANNEXE E.....	84
	ANNEXE E - TAUX DE RÉPONSE PAR RÉGION SOCIO-SANITAIRE ET PAR BASE DE SONDAGE DE L'ESCC 2009	85

1. INTRODUCTION

L'Enquête sur la santé dans les collectivités canadiennes (ESCC) est une enquête transversale qui vise à recueillir des renseignements sur l'état de santé, l'utilisation des services de santé et les déterminants de la santé de la population canadienne. Elle est réalisée auprès d'un grand échantillon de répondants et conçue pour fournir des estimations fiables à l'échelle de la région sociosanitaire. En 2007, des changements importants ont été apportés à la conception de l'ESCC. Les données sont dorénavant recueillies sur une base continue et leur diffusion est prévue chaque année, plutôt que chaque deux ans comme c'était le cas avant 2007. Les objectifs de l'enquête ont été révisés. Ces objectifs sont :

- soutenir les programmes de surveillance en santé en produisant des données sur la santé à l'échelle nationale, provinciale et infraprovinciale;
- offrir une source unique de renseignements pour la recherche sur la santé de petites populations et sur des caractéristiques rares;
- diffuser de l'information facilement accessible à une communauté diversifiée d'utilisateurs dans un temps opportun; et
- proposer un instrument d'enquête flexible qui inclut une option de réponse rapide pour répondre à des questions émergentes liées à la santé de la population.

Les autres changements qui ont découlé du remaniement sont détaillés à la section 3.

Les données de l'ESCC sont toujours recueillies auprès de personnes âgées de 12 ans et plus vivant dans des logements privés dans 121 régions sociosanitaires couvrant toutes les provinces et les territoires. Sont exclues de la base de sondage les personnes vivant sur les réserves indiennes et les terres de la Couronne, les résidents des établissements, les membres à temps plein des Forces canadiennes et les personnes vivant dans certaines régions éloignées. L'ESCC couvre environ 98 % de la population canadienne âgée de 12 ans et plus.

Le présent document a pour but de faciliter la manipulation des fichiers de microdonnées de l'ESCC et de décrire la méthodologie utilisée. L'enquête produit 3 types de fichiers de microdonnées : des fichiers maîtres, des fichiers de partage et des fichiers de microdonnées à grande diffusion (FMGD). Les caractéristiques de chacun de ces fichiers sont présentées dans ce guide. Les données du FMGD, dont la diffusion est bisannuelle, couvrent deux années. Le prochain FMGD, qui contiendra les données recueillies pour les années 2009 et 2010, sera diffusé en 2011.

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

Service d'aide aux utilisateurs des produits électroniques : 1 (800) 949-9491

Totalisations spéciales ou renseignements généraux sur les données :

Services personnalisés à la clientèle, Division de la statistique de la santé : (613) 951-1746

Courriel : hd-ds@statcan.gc.ca

Renseignements sur le télé-accès :

(613) 951-1746

Courriel :

cchs-escc@statcan.gc.ca

Télécopieur :

(613) 951-0792

2. HISTORIQUE

En 1991, le Groupe de travail national sur l'information en matière de santé a relevé plusieurs questions et problèmes posés par le système d'information sur la santé. Selon ses membres, les données étaient fragmentées, elles étaient incomplètes, elles ne pouvaient être partagées facilement et elles n'étaient pas analysées aussi pleinement que possible; en outre, les résultats des études réalisées n'atteignaient pas de façon régulière la population canadienne¹.

Pour résoudre ces problèmes, l'Institut canadien d'information sur la santé (ICIS), Statistique Canada et Santé Canada ont conjugué leurs efforts en vue de créer un Carnet de route de l'information sur la santé. L'Enquête sur la santé dans les collectivités canadiennes (ESCC) a été conçue à partir de ce mandat. Le format, le contenu et les objectifs de l'enquête ont été définis après avoir mené des consultations approfondies auprès de spécialistes et d'intervenants fédéraux, provinciaux et de régions sociosanitaires communautaires en vue de déterminer leurs exigences en matière de données².

Afin de remplir les nombreux besoins en données, le cycle de collecte des données de l'ESCC s'étendait sur deux années. Jusqu'au remaniement de 2007, la première année du cycle, indiquée par la notation « .1 », correspondait à une enquête générale sur la santé de la population conçue pour fournir des estimations fiables à l'échelle de la région sociosanitaire. La deuxième année du cycle, représentée par la notation « .2 », avait un plus petit échantillon et était conçue pour fournir des données à l'échelle provinciale sur des sujets particuliers ayant trait à la santé.

Nouvelles désignations des Cycles .1 et .2

À partir de 2007, la composante régionale du programme de l'ESCC est dorénavant collectée sur une base continue. Pour éviter toute confusion avec les enquêtes de santé thématiques, les deux composantes ont cessé d'utiliser les notations « .1 » et « .2 » pour se distinguer entre elles. Dorénavant, on désigne les cycles x.1 de l'ESCC comme « la composante annuelle » de l'ESCC. Le titre au long est « Enquête sur la santé dans les collectivités canadiennes – Composante annuelle, 2009 » et le titre abrégé est simplement « ESCC – 2009 ». La composante thématique de l'enquête demeure, quant à elle, inchangée. Elle continuera d'explorer plus en profondeur des sujets ou des populations plus précises. On la désignera par le nom de l'enquête suivi du sujet des thèmes couverts par chacune des enquêtes (par exemple, « Enquête sur la santé dans les collectivités canadiennes portant sur le vieillissement en santé » ou « ESCC – Vieillissement en santé »).

¹ 1999. Carnet de route de l'information sur la santé — Répondre aux besoins, Santé Canada, Statistique Canada. p. 3.

² 1999. Initiative du carnet de route ... Lancer le processus. Institut canadien d'information sur la santé/Statistique Canada. ISBN 1-895581-70-2. p. 19.

3. REMANIEMENT DE L'ESCC EN 2007

Jusqu'en 2005, les données de l'ESCC étaient recueillies à chaque deux ans sur une période annuelle et diffusées aux deux ans, environ 6 mois après la fin de la collecte. Le remaniement de l'ESCC de 2007 visait deux principaux points : mieux répondre aux besoins des partenaires qui désiraient augmenter le contenu de l'enquête et la fréquence des diffusions de données et assurer une meilleure utilisation des ressources opérationnelles. C'est ainsi que des changements à la conception de l'ESCC ont été proposés de manière à augmenter l'efficacité et la flexibilité de l'enquête par une collecte des données sur une base continue.

Des consultations approfondies ont été menées d'un bout à l'autre du Canada auprès de spécialistes et intervenants fédéraux, provinciaux et de régions sociosanitaires. Elles visaient à recueillir les commentaires sur les changements proposés ainsi qu'à recueillir des renseignements détaillés sur les besoins en données et produits de données des divers partenaires.

Les principaux changements qui ont découlé du remaniement de l'ESCC sont les suivants:

- Dans le passé, les données de l'ESCC étaient collectées auprès de 130 000 répondants sur une période de 12 mois. Dorénavant, la collecte de données se déroule sur une base continue. L'échantillon, qui préserve la même taille, est réparti en 12 période de collecte de 2 mois chacune. Chacune des périodes de collecte est représentative de la population vivant dans les 10 provinces canadiennes au cours des deux mois. Pour des raisons opérationnelles, l'échantillon dans les territoires est représentatif de leur population après 12 mois.
- La composante de contenu commun a été scindée en trois : le contenu commun annuel (auparavant le « contenu de base »), le contenu commun un an qui couvre un an de collecte et le contenu commun deux ans qui couvre deux ans de collecte (auparavant le « contenu thématique »). Demandé pour une année, le contenu commun un an est réintroduit tous les deux ou quatre ans. Demandé pour deux années, le contenu commun deux ans, couvrant deux années, est réintroduit tous les quatre ans. L'un et l'autre ont été créés pour tirer profit de l'approche de collecte continue. La durée de collecte des données de cette composante peut être adaptée selon la prévalence des estimations et le niveau géographique désirés. Le contenu commun annuel demeure relativement stable dans le temps. À la discrétion des provinces et des régions, le contenu optionnel peut aussi être modifié sur une base annuelle, plutôt que tous les deux ans.
- Les changements à la collecte et au contenu ont inévitablement un impact sur la stratégie de diffusion. Dans le passé, les données étaient diffusées à tous les deux ans. Depuis 2008, les données de l'ESCC sont diffusées sur une base annuelle. À tous les deux ans, un fichier regroupant l'échantillon des deux années (taille de 130 000) est également produit. En plus des ces fichiers réguliers, d'autres fichiers spéciaux seront rendus disponibles lorsque du contenu supplémentaire aura été recueilli pendant des périodes de collecte qui ne correspondent pas aux périodes annuelles standard, c'est-à-dire de janvier à décembre.
- La collecte annuelle de données est divisée en 6 périodes de 2 mois. Contrairement à ce qui se faisait avant, ces périodes ne se chevauchent plus, ce qui permet d'assurer une

surveillance plus efficace de la collecte et d'effectuer au besoin des modifications à l'interface de collecte à tous les deux mois.

4. CONTENU DE L'ESCC

En plus de données sociodémographiques et administratives, le contenu de l'ESCC comporte trois composantes qui répondent chacune à des besoins différents : la composante de contenu commun inclut le contenu commun annuel, le contenu d'un an et le contenu de deux ans, la composante de contenu optionnel et la composante de réponse rapide. L'annexe A présente la liste des modules qui ont été inclus dans le questionnaire de 2009 selon leur composante.

La durée moyenne d'une entrevue de l'ESCC est estimée entre 40 et 45 minutes.

Tableau 4.1 Durée de l'enquête selon les composantes

4.1 Contenu commun

La composante de contenu commun de l'ESCC comprend des questions qui sont demandées aux répondants de toutes les provinces et territoires (à moins d'exception Cette composante est scindée en trois : le contenu commun annuel, le contenu commun un an et le contenu commun deux ans.

Le contenu commun annuel inclut des questions qui sont demandées à tous les répondants. Ces questions demeureront relativement stables dans le questionnaire pour une période d'environ six ans, à moins qu'un enjeu majeur soit soulevé concernant la qualité des données.

Le contenu commun un an et deux ans (auparavant appelé « contenu thématique ») comprennent des questions se rapportant à un sujet particulier. Combinés, ils comptent pour 10 minutes de l'entrevue. Les modules comportant ce type de contenu pourraient être réintroduits dans l'enquête tous les deux, quatre ou six ans, au besoin. Ceci permet une meilleure planification du contenu de l'ESCC à moyen terme.

Certains modules inclus dans le contenu commun un an peuvent être demandés à un sous-échantillon de répondants si l'objectif de ces questions est de fournir des estimations fiables à l'échelle nationale ou provinciale, plutôt qu'à l'échelle de la région sociosanitaire. Cette approche est utilisée dans le but de minimiser le fardeau de réponse et les coûts qui s'y rattachent.

4.2 Contenu optionnel

La composante de contenu optionnel offre aux régions sociosanitaires de choisir du contenu qui répond aux priorités provinciales et régionales en matière de santé publique. Le contenu optionnel est sélectionné à partir d'une longue liste de modules disponibles pour inclusion dans l'ESCC. Les modules de contenu choisis par une région ne sont demandés qu'aux résidents des régions qui ont sélectionné ces modules. En réalité, depuis 2005 (cycle 3.1), les régions et les provinces ont choisi de coordonner la sélection du contenu optionnel de manière à uniformiser la sélection des modules optionnels à l'échelle provinciale. Le contenu optionnel peut varier d'une année à l'autre selon les besoins et doit être revu chaque deux ans.

Il convient de mentionner que contrairement au contenu commun, les données provenant des modules de contenu optionnel ne peuvent être généralisées à l'échelle du Canada³.

L'annexe B présente les résultats de la sélection du contenu optionnel pour l'année en cours selon la province de résidence.

4.3 Contenu de réponse rapide

La composante de réponse rapide est offerte contre recouvrement des coûts aux organisations désirant obtenir des estimations nationales sur un sujet émergent ou particulier lié à la santé de la population. La réponse rapide est formée d'un maximum de deux minutes de contenu. Les questions apparaissent au questionnaire pendant une seule période de collecte (deux mois) et sont demandées à tous les répondants de l'ESCC au cours de cette période.

4.4 Contenu des fichiers de microdonnées

Différents fichiers de données sont produits à partir des données de l'enquête :

- période de référence d'un an;
- période de référence de deux ans;
- sous-échantillon d'un an.

Le tableau 4.2 fournit des précisions concernant la disponibilité du contenu dans les fichiers de, 2009.

Fichiers à période de référence d'un an

³ À moins qu'un module optionnel soit sélectionné par l'ensemble des régions sociosanitaires du Canada au cours d'une même période de collecte, ce qui n'est jamais produit à ce jour.

L'enquête produit des fichiers tous les ans.. En juin 2010, un fichier dont la période de référence est 2009 a été diffusé. Il contient les répondants de la collecte de 2009 ainsi que les variables du contenu commun annuel, celles du contenu commun un an, celles du contenu commun de deux ans et du contenu optionnel.

Fichier à période de référence de deux ans

Chaque deux ans, un fichier combinant les deux années les plus récentes de données est également diffusé. Le dernier fichier combiné, diffusé en 2009, contenait des données pour 2007 et 2008. Un fichier combiné sera diffusé en 2011 et inclura les données des années de référence 2009 et 2010.

Le fichier de deux ans inclut tous les répondants ainsi que l'ensemble des questions qui sont demeurées dans l'enquête au cours de cette période. À moins d'exception, il s'agit de la composante des questions du contenu commun annuel, deux ans et le contenu optionnel choisi pour deux ans. Le contenu commun un an et le contenu optionnel choisi pour une année seulement ne sont pas disponibles dans le fichier de données de deux ans.

Fichiers de sous-échantillons d'un an

Les modules collectés auprès d'un sous-échantillon de la population continueront d'être diffusés dans des fichiers séparés. Ces derniers incluent le contenu commun annuel et un an collectés auprès d'un sous-échantillon de répondants.

Tableau 4.2 Composantes de contenu comprises dans les fichiers de données de 2009 et 2010

¹ Le commun un an de 2009 était composé de 2 modules (Accès aux services de santé et Temps d'attentes) qui ont été posés à un sous-échantillon de répondants.

² Le contenu commun un an de 2010 comprendra un groupe de modules reliés à la détection des maladies chroniques.

³ Le contenu optionnel qui sera inclut dans le fichier de 2009-2010 aura été demandé aux répondants d'une province pendant les deux années. Autrement, il sera inclut uniquement dans le fichier pour l'année pendant laquelle il a été recueilli. Il est à noter que si un module de contenu commun un an est choisi dans le contenu optionnel d'une juridiction au cours de la deuxième année, le module sera inclut dans le fichier de deux ans et traité comme contenu optionnel.

5. PLAN D'ÉCHANTILLONNAGE

5.1 *Population cible*

L'ESCC vise la population de 12 ans et plus vivant à domicile et résidant dans les dix provinces et trois territoires. Sont exclues du champ de l'enquête les personnes vivant sur les réserves indiennes et les terres de la Couronne, les résidents des établissements, les membres à temps plein des Forces canadiennes et les personnes vivant dans certaines régions éloignées. L'ESCC couvre environ 98 % de la population canadienne de 12 ans et plus.

5.2 Régions sociosanitaires

À des fins administratives, chaque province est divisée en plusieurs régions sociosanitaires (RS) et chaque territoire est considéré comme formant une RS unique. En collaboration avec les provinces, Statistique Canada est parfois appelé à modifier les limites de certaines RS afin qu'elles correspondent aux données géographiques du recensement ou encore afin de mieux tenir compte des besoins en données sur la santé selon de nouvelles bornes géographiques. Pour l'ESCC 2009, des données ont été recueillies pour 118 RS dans les 10 provinces, ainsi que pour une RS par territoire, pour un total de 121 RS (Annexe C).

5.3 Taille et répartition de l'échantillon

Afin de produire des estimations fiables pour chaque RS et compte tenu du budget accordé pour l'ESCC, il a été établi que l'enquête devait être réalisée auprès d'un échantillon d'environ 130 000 personnes sur une période de 2 ans. La production d'estimations fiables pour chaque RS était l'objectif principal, mais la qualité des estimations pour certaines caractéristiques importantes pour les provinces a été jugée importante également. Par conséquent, la stratégie de répartition de l'échantillon, qui comporte trois étapes, accorde une importance plus ou moins égale aux RS et aux provinces. A la première étape, on impose une taille minimum de 500 répondants par RS, considérée comme étant le minimum requis afin d'obtenir un niveau de qualité de données raisonnable. Par contre, pour des raisons de fardeau de réponse, on impose une fraction de sondage maximum de 1 sur 20 logements afin d'éviter d'échantillonner trop de logements dans des petites régions qui sont aussi interpellées par d'autres enquêtes. Notons que très peu de RS ont une taille de moins de 500 due à la limite de la fraction de sondage. Cette première étape a réparti au total 60,350 unités. La deuxième étape consiste à répartir le reste de l'échantillon disponible par province en suivant une répartition proportionnelle à la taille de la population par province. La taille totale de l'échantillon par province est donc la somme des tailles établies aux deux premières étapes. Il est à noter que la stratégie de répartition de l'échantillon a été utilisée pour l'ESCC au cycle 3.1 et que les tailles alors établies ont restées sensiblement les mêmes depuis. L'échantillon est ensuite divisé également pour chacune des 2 années de collecte. Le tableau 5.1 donne la taille d'échantillon annuel pour 2009.

Tableau 5.1 Nombre de régions sociosanitaires et tailles visées d'échantillon selon la province/territoire pour l'ESCC 2009

Province	Nombre de RS	Taille totale de l'échantillon (visée)
Terre-Neuve-et-Labrador	4	2 005
Île-du-Prince-Édouard	3	1 001
Nouvelle-Écosse	6	2 520
Nouveau-Brunswick	7	2 575
Québec	16	12 144
Ontario ¹	36	22 207
Manitoba	10	3 750
Saskatchewan	11	3 860
Alberta	9	6 100
Colombie-Britannique	16	8 050
Yukon	1	600
Territoires du Nord-Ouest	1	600
Nunavut	1	350
Canada	121	65 762

¹La taille de l'échantillon pour l'Ontario inclut l'achat d'échantillons supplémentaires par réseau local d'intégration des services de santé (RLISS). La taille initiale pour l'Ontario avant l'achat était de 20 880 unités (se référer à la section 5.7 pour plus de détails)

À la troisième étape, l'échantillon provincial a été réparti entre les RS proportionnellement à la racine carrée de la population estimée de la RS. Cette stratégie en trois étapes permet d'obtenir un échantillon suffisant pour chaque RS, sans perturber considérablement la répartition proportionnelle à la taille par province.

Il convient de souligner que les trois territoires, qui ont été traités séparément, n'étaient pas visés par la stratégie susmentionnée de répartition de l'échantillon. Au total, pour 2009, 600 unités d'échantillonnage ont été attribuées au Yukon, 600 aux Territoires du Nord-Ouest et 350 au Nunavut. Ces tailles sont déterminées selon le budget disponible. La répartition de l'échantillon dans les territoires est faite proportionnellement à la taille de la population dans les strates. Les strates utilisées sont les mêmes que celles définies par l'Enquête sur la population active (EPA), qui correspondent à des regroupements de communautés. Pour plus de détails, se reporter à la section 5.4.1.

L'échantillon est ensuite divisé également en deux, de sorte à obtenir les mêmes tailles d'échantillon entre l'échantillon de la base aréolaire et l'échantillon de la base liste pour chaque RS⁴ tel qu'élaboré dans la section suivante. Mentionnons finalement que la taille des échantillons tirés de chaque base a été augmentée avant la collecte des données afin de tenir compte des unités hors

⁴ A l'exception de 2 régions sociosanitaires utilisant la base de sondage à composition aléatoire (CA) seulement (section 5.4.3) et les 3 territoires utilisant seulement la base aréolaire et la base de sondage à composition aléatoire (sections 5.4.1 et 5.4.3).

du champ de l'enquête et du taux de non-réponse anticipés à partir des taux obtenus lors des cycles précédents de l'ESCC. Les tailles d'échantillons par RS et par base de sondage sont données à l'annexe D.

5.4 Bases de sondage et stratégies d'échantillonnage des ménages

L'échantillon de ménages de l'ESCC 2009 a été sélectionné à partir de trois bases de sondage : 49 % de l'échantillon de ménages provenait d'une base de sondage aréolaire, 50 % provenait d'une base liste de numéros de téléphone et 1 % provenait d'une base de sondage à composition aléatoire (CA).

5.4.1 Échantillonnage des ménages à partir de la base aréolaire

La base aréolaire conçue pour l'Enquête sur la population active (EPA) du Canada a servi de base de sondage principale pour l'ESCC. 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'échantillonnage⁵. À la première étape, des strates homogènes sont formées et des échantillons indépendants de grappes sont sélectionnés dans chaque strate. À la deuxième étape, une liste de logements pour chaque grappe est créée, puis des logements sont sélectionnés dans chaque liste.

Pour les besoins du plan d'échantillonnage de l'EPA, chaque province est répartie en trois catégories de région, soit les grands centres urbains, les villes et les régions rurales. Des strates géographiques ou socioéconomiques sont formées à l'intérieur de chaque grand centre urbain. Dans les strates, des grappes sont formées par regroupement de 150 à 250 logements. Dans certains centres urbains, des strates distinctes sont créées pour les appartements ou pour les aires de diffusion (AD) du recensement pour cibler les ménages à haut revenu, les immigrants et les autochtones. Dans chaque strate, six grappes ou bâtiments d'habitation (parfois 12 ou 18 appartements) sont sélectionnés par une méthode d'échantillonnage aléatoire avec probabilité proportionnelle à la taille (PPT), cette dernière correspondant au nombre de ménages. Le nombre 6 est utilisé pour l'ensemble du plan d'échantillonnage afin de permettre le renouvellement mensuel d'un sixième de l'échantillon de l'EPA.

Les autres villes et régions rurales de chaque province sont stratifiées, en premier lieu, en fonction de données géographiques, puis selon les caractéristiques socioéconomiques. Dans la plupart des strates, six grappes (habituellement des AD du recensement) sont sélectionnées par la méthode PPT. Certains centres urbains isolés géographiquement sont couverts par un plan de sondage à trois degrés. Ce type de plan de sondage est utilisé au Québec, en Ontario, en Alberta et en Colombie-Britannique.

Une fois la liste des nouvelles grappes établie, l'échantillon est tiré par échantillonnage systématique des logements. La taille de chaque échantillon systématique est appelée le 'rendement'. Le tableau 5.2 donne un aperçu des catégories d'unités primaires d'échantillonnage (UPE) utilisées dans l'échantillon de l'EPA et le rendement prévu par échantillon systématique. Comme les taux d'échantillonnage sont prédéterminés, il existe souvent un écart entre la taille

⁵ Statistique Canada (1998). *Méthodologie de l'enquête sur la population active du Canada*. Statistique Canada numéro 71-526-XPB au catalogue.

prévue d'échantillon et les chiffres obtenus. Par exemple, le rendement de l'échantillon est parfois excessif. Cette situation peut se présenter dans des secteurs où le nombre de logements a augmenté à la suite de nouveaux projets de construction. Pour réduire le coût de la collecte des données, la production excessive est corrigée en éliminant, dès le départ, une partie des unités sélectionnées. Cette modification est considérée lors de la pondération.

Tableau 5.2 Unité primaire d'échantillonnage, taille et rendement

Région	Unité primaire d'échantillonnage (UPE)	Taille (ménages par UPE)	Rendement (ménages échantillonnés)
Toronto, Montréal, Vancouver	Grappe	150 à 250	6
Autres villes	Grappe	150 à 250	8
La plupart des régions rurales/petits centres urbains	Grappe	100 à 250	10

Afin de répondre aux exigences particulières de l'ESCC, certaines modifications ont dû être apportées à cette stratégie d'échantillonnage. Pour obtenir un échantillon annuel de 33 000 répondants pour l'ESCC 2009, il a fallu sélectionner près de 48 000 logements de la base aréolaire afin de tenir compte des logements vacants et des ménages non-répondants. Chaque mois, le plan d'échantillonnage de l'EPA fournit environ 60 000 logements répartis entre les diverses régions économiques des dix provinces alors que, pour l'ESCC 2009, il fallait obtenir un échantillon d'environ 48 000 logements réparti par RS dont les limites géographiques différaient de celles des régions économiques de l'EPA. Globalement, l'ESCC 2009 nécessitait la sélection d'un nombre inférieur de logements que produit le mécanisme de sélection de l'EPA, ce qui correspond à un *facteur de redressement* de 0,80 (48 000/60 000). Toutefois, comme ce facteur de redressement variait de 0,3 à 3,0 au niveau des RS, certains ajustements ont été nécessaires.

Les modifications apportées au processus de sélection dans les régions variaient selon le facteur de redressement. Pour les RS où le facteur était égal ou inférieur à 1, le nombre d'UPE choisies a été réduit si nécessaire. Par exemple, si le facteur était de 0,5, alors seulement 3 UPE ont été choisies dans chaque strate de l'EPA au lieu du nombre habituel de 6 UPE. Pour les RS avec un facteur supérieur à 1 mais égal ou inférieur à 2, le processus d'échantillonnage des logements à l'intérieur d'une UPE a été répété pour un sous-ensemble des UPE sélectionnées appartenant à la RS en question. Par exemple, si le facteur était de 1,6, alors la sélection des logements a été répétée dans 4 des 6 UPE pour chaque strate dans la RS. Lorsque la répétition de la sélection de logements à l'intérieur d'une UPE était nécessaire mais qu'aucun autre logement n'était disponible dans cette UPE, alors une autre UPE a été choisie. Lorsque le facteur était supérieur à 2, le processus d'échantillonnage des logements a été répété à l'intérieur d'autres UPE appartenant à la RS en question⁶.

⁶ Pour réduire les coûts de listage, le processus de sélection a été répété jusqu'à 3 fois dans les UPE déjà choisies, pour certaines régions urbaines seulement. Il s'agit par contre de cas d'exception.

Finalement, lorsque le nombre de logements disponibles dans les UPE sélectionnées était supérieur au nombre de logements requis, un sous-échantillon a été tiré. Ce processus est appelé la 'stabilisation'.

Échantillonnage des ménages à partir de la base aréolaire dans les trois territoires

Pour des raisons opérationnelles, le plan d'échantillonnage à partir de la base aréolaire de l'EPA pour les trois territoires est différent. Pour chaque territoire, les collectivités (municipalités) faisant partie du champ de l'enquête sont groupées en strates en fonction de diverses caractéristiques (population, données géographiques, proportion d'Inuit et/ou d'Autochtones et revenu médian du ménage). L'EPA a défini 5 strates pour le Yukon, 10 pour les Territoires du Nord-Ouest et 6 pour le Nunavut. Le premier degré d'échantillonnage consiste à sélectionner aléatoirement une collectivité avec probabilité proportionnelle à la taille de la population dans chaque strate. Puis, à l'intérieur de chaque collectivité, une stratégie d'échantillonnage des ménages est mise en place de façon identique à celle décrite plus haut. L'ESCC a sélectionné son échantillon à partir des mêmes collectivités sélectionnées par l'EPA tout en s'assurant de sélectionner des logements différents. Si trop ou pas assez de logements étaient disponibles pour une collectivité à l'intérieur d'une strate, l'EPA a choisi une autre collectivité pour l'ESCC.

Il convient de mentionner que la base de sondage de l'ESCC 2009 couvrait 90 % des ménages privés du Yukon, 97 % de ceux des Territoires du Nord-Ouest et 71 % de ceux du Nunavut⁷.

5.4.2 Échantillonnage des ménages à partir de la base liste de numéros de téléphone

À l'exception de 5 RS (les 2 RS provenant de la base CA et les trois territoires), la base liste de numéros de téléphone a été utilisée dans toutes les régions pour compléter la base aréolaire. La base liste est l'annuaire téléphonique du Canada, une base administrative externe de numéros de téléphone contenant les noms, les adresses et les numéros de téléphone répertoriés dans les annuaires de téléphone du Canada et qui est mise à jour à tous les 6 mois. Elle a été stratifiée par RS en utilisant un fichier de conversion de codes postaux de sorte à pouvoir associer une RS à chaque numéro de téléphone. Dans chaque RS, un échantillon de numéros de téléphone a été tiré par échantillonnage aléatoire simple. Comme pour la base de sondage à CA, des numéros de téléphone supplémentaires ont été sélectionnés pour tenir compte des numéros hors service ou hors du champ d'observation.

Il importe de souligner que le sous-dénombrement de la base liste de numéros de téléphone est plus important que celui de la base de sondage à CA, car les numéros non publiés n'ont aucune chance d'être sélectionnés. Néanmoins, comme la base liste est toujours utilisée en complément de la base aréolaire, l'effet du sous-dénombrement dû à l'utilisation de la base liste de numéros de téléphone est minimal et est corrigé lors de la pondération.

⁷ Au Nunavut, pour des raisons d'ordre opérationnel inhérentes aux collectivités éloignées, seulement les 10 communautés les plus grandes sont couvertes par l'enquête soient Iqaluit, Cambridge Bay, Baker Lake, Arviat, Rankin Inlet, Kugluktuk, Pond Inlet, Cape Dorset, Pangnirtung et Igloodik.

5.4.3 Échantillonnage des ménages à partir de la base de sondage à CA de numéros de téléphone

Dans 4 RS, un échantillon de numéros de téléphone provenant de la base de sondage à composition aléatoire (CA) a été utilisé pour sélectionner un échantillon de ménages. L'échantillonnage de ménages à partir de la base à CA a été réalisé selon la méthode d'élimination des banques non valides (EBNV) adoptée par l'Enquête sociale générale⁸. Une banque de cent numéros (c'est-à-dire les huit premiers chiffres d'un numéro de téléphone à 10 chiffres) est considérée comme non valide si elle ne contient aucun numéro de téléphone résidentiel. Au départ, la base de sondage comprend la liste de toutes les banques valides de cent numéros et celles qui ne sont pas valides sont éliminées de la base de sondage à mesure qu'on les repère. Il convient de souligner que ces banques de cent numéros ne sont éliminées de la base de sondage que lorsque l'on possède des preuves qu'elles ne sont pas valides provenant de sources diverses multiples. En l'absence de renseignements, la banque est retenue dans la base de sondage. Pour éliminer les banques non valides, on s'est servi de l'annuaire du téléphone, ainsi que de divers fichiers administratifs internes.

D'après les renseignements géographiques disponibles (codes postaux), les banques de cent numéros retenues dans la base de sondage ont été regroupées pour créer des strates CA englobant, de façon aussi exacte que possible, les régions sociosanitaires. À l'intérieur de chaque strate CA, une banque de cent numéros a été choisie au hasard et un numéro compris entre 00 et 99 a été généré aléatoirement afin de créer un numéro de téléphone complet à 10 chiffres. Cette méthode a été répétée jusqu'à ce que l'on ait atteint le nombre requis de numéros de téléphone pour la strate CA. Comme fréquemment, le numéro obtenu n'est pas en service ou est hors du champ d'observation, il faut générer un grand nombre de numéros supplémentaires pour atteindre la taille visée d'échantillon. Ce taux de réussite diffère selon la région. Dans le cas de l'ESCC, il variait de 25 % à 50% parmi les 4 RS qui y ont eu recours.

5.5 Répartition de l'échantillon par période de collecte des données

Afin d'équilibrer la charge de travail des intervieweurs et de réduire au minimum les effets saisonniers éventuels sur les estimations de caractéristiques importantes telle que l'activité physique, dans chaque RS, l'échantillon initial de logements/numéros de téléphone a été réparti au hasard, de façon égale pour chaque période de collecte de 2 mois.

Pour l'échantillon de la base aréolaire, chaque UPE sélectionnée dans chaque RS a été répartie au hasard à une période de collecte en tenant compte de plusieurs contraintes reliées aux opérations sur le terrain ou encore à la pondération tout en conservant une taille égale par période. Par exemple, on s'est assuré d'avoir un échantillon représentatif de la population canadienne aux 6 mois en s'assurant d'avoir un échantillon de logements couvrant toutes les strates de l'EPA sur cette période.

Pour les listes des numéros de téléphone, des échantillons indépendants ont été sélectionnés à chaque période de collecte. Cette stratégie permet d'assurer que chaque échantillon soit

⁸ Norris, D.A., Paton, D.G. (1991), L'Enquête sociale générale canadienne: bilan des cinq premières années. *Techniques d'enquête* (Statistique Canada, Catalogue 12-001); 17, pp. 245-260.

représentatif de la population canadienne faisant partie du champ d'observation de l'enquête à chaque 2 mois.

5.6 Échantillonnage des personnes interviewées

Comme pour les cycles précédents, la sélection des répondants a été conçue de façon à ce que les jeunes (de 12 à 19 ans) soient surreprésentés dans l'échantillon. La stratégie d'échantillonnage adoptée a tenu compte des besoins des utilisateurs de données, du coût, de l'efficacité du plan d'échantillonnage, du fardeau de réponse et des contraintes opérationnelles. Une personne est sélectionnée par ménage en utilisant diverses probabilités de sélection variant selon l'âge et selon la composition du ménage. Les probabilités choisies font suite à des résultats de simulations reposant sur divers paramètres dans le but de déterminer l'approche optimale sans générer de poids d'échantillonnage extrêmes en bout de ligne.

Le tableau 5.3 donne les facteurs multiplicatifs de poids de sélection utilisés pour déterminer les probabilités de sélection des personnes dans les ménages échantillonnés, par groupe d'âge. Par exemple, pour un ménage de trois personnes formé de deux adultes âgés entre 45 et 64 et d'un jeune de 15 ans, le jeune de 15 ans aurait 6.5 fois plus de chances d'être sélectionné que chacun des adultes. Afin d'éviter l'obtention de poids extrêmes, il y a une exception à la règle: si la taille du ménage est plus grande ou égale à 5 ou si le nombre de personnes âgées de 12 à 19 ans est plus grand ou égal à 3, alors le facteur multiplicatif du poids de sélection est égal à 1 pour toutes les personnes du ménage. Dans ce cas, toutes les personnes du ménage ont la même probabilité d'être sélectionnées.

Tableau 5.3 Facteurs multiplicatifs du poids de sélection pour la stratégie d'échantillonnage au niveau de la personne, par âge

Âge	Facteur multiplicatif
12-19	6.5
20-24	1.5
25-34	1.5
35-44	1.5
45-64	1.5
65+	1.5

5.7 Achat d'échantillon pour l'Ontario

La province de l'Ontario a demandé une augmentation de l'échantillon afin de produire des estimations au niveau de la géographie des réseaux locaux d'intégration des services de santé (RLISS). L'Ontario compte 14 RLISS. L'échantillon de l'ESCC a été augmenté de sorte à obtenir une taille minimum de 2000 par RLISS sur une période de 2 ans. Comme les limites des RS et RLISS s'entrecoupent, le niveau de stratification utilisé a été le croisement RS-RLISS. Les tailles d'échantillon préalablement allouées par RS ont donc été conservées. Dans les cas où la répartition par RS n'a pas permis d'atteindre des tailles de 2000 par RLISS, l'échantillon a été augmenté en conséquence et réparti proportionnellement à la taille de la population dans le croisement RS-RLISS. Le tableau 5.4 donne la taille d'échantillon de répondants visés par RLISS pour l'année 2009.

Tableau 5.4 Tailles visées de répondants par réseau local d'intégration des services de santé (RLISS) dans l'ESCC 2009

RLISS	Taille visée de répondants
01-Erie St. Clair	1 550
02-South West	2 561
03-Waterloo Wellington	1 242
04-Hamilton Niagara Haldimand Brant	2 597
05-Central West	1 069
06-Mississauga Halton	1 115
07-Toronto Central	1 081
08-Central	1 411
09-Central East	2 108
10-South East	1 313
11-Champlain	2 057
12-North Simcoe Muskoka	1 050
13-North East	1 990
14-North West	1 063
Ontario	22 207

La taille totale de l'échantillon des croisements RS-RLISS a ensuite été répartie également entre la base liste et la base aréolaire. Les procédures normales de sélection de l'échantillon dans chaque base ont été appliquées sur l'échantillon total. L'échantillon supplémentaire fait partie intégrante de l'échantillon de l'ESCC. Les tailles d'échantillon par réseau local d'intégration des services de santé et par base de sondage sont données à l'annexe D.

5.8 Sous-échantillonnage pour l'Enquête sur l'accès aux services de santé (EASS)

Un sous-échantillon de l'échantillon de l'ESCC a été tiré afin d'obtenir de l'information supplémentaire sur l'accès aux services de soins de santé. L'enquête couvre la même population que l'ESCC à l'exception des territoires et des personnes âgées de moins de 15 ans.

Le budget alloué à ce sous-échantillon était similaire à l'enquête précédente soit près de 49 000 répondants, permettant de produire des estimations fiables au niveau provincial. De plus, les mêmes tailles d'échantillon étaient visées par province. Cependant, la taille d'échantillon à l'Île-du-Prince-Édouard n'a pas été gonflée comme en 2007; seul l'échantillon de l'ESCC de 1 001 unités a été utilisé. Voici donc les tailles finales pour le sous-échantillon de l'EASS sélectionné à partir de l'ESCC 2009.

Table 5.5 Tailles des échantillons de l'ESCC 2009 et de l'EASS 2009, par province

Province	Taille d'échantillon	
	ESCC 2009	EASS 2009
Terre-Neuve-et-Labrador	2 005	2 005
Île-du-Prince-Édouard	1 001	1 001
Nouvelle-Écosse	2 520	2 520
Nouveau-Brunswick	2 575	2 575
Québec	12 144	4 600
Ontario	22 207	22 207
Manitoba	3 750	3 200
Saskatchewan	3 860	3 200
Alberta	6 100	3 600
Colombie-Britannique	8 050	4 000
CANADA	64 212	48 908

Une fois la taille définie par province, l'échantillon a été réparti par RS proportionnellement à la taille de la population par RS, permettant ainsi une meilleure répartition de l'échantillon par province tout en tenant compte de la stratification de l'ESCC par RS. Pour les provinces où la taille d'échantillon par RS était insuffisante, une répartition de puissance avec puissance inférieure à 1 a dû être utilisée. Une puissance de 0,9 a été utilisée en Alberta et en Colombie-Britannique, alors qu'une puissance de 0,55 a dû être utilisée au Manitoba et en Saskatchewan, rendant le plan moins optimal. Pour les autres provinces, aucune répartition par RS n'était nécessaire puisque l'échantillon complet de l'ESCC a été utilisé.

Finalement, la taille a été répartie également entre la base liste et la base aréolaire. La taille a également été gonflée afin de tenir compte des unités hors du champ de l'enquête, et du taux prévu de non-réponse. Lorsque possible, la taille a également été gonflée afin de tenir compte de la population non-couverte par l'EASS (les 12-14 ans) et de tenir compte des entrevues par personne interposée qui ne sont pas acceptées dans l'EASS. Les tailles d'échantillon par province et par base de sondage du sous-échantillon de l'EASS 2009 sont données à l'annexe D.

La sélection de l'échantillon s'est faite de façon indépendante à chaque période de collecte à partir des échantillons de l'ESCC 2009. Un sous-échantillon de logements ou de numéros de téléphone de même taille a été tiré aléatoirement dans chaque RS à chaque période de collecte.

6. COLLECTE DES DONNÉES

6.1 *Interviews assistées par ordinateur*

Entre janvier et décembre 2009, un total de 61 679 interviews valables assistées par ordinateur (IAO) ont été effectuées. Environ la moitié ont eu lieu au moyen de la méthode de l'interview sur place assistée par ordinateur (IPAO), l'autre moitié ayant consisté en des interviews téléphoniques assistées par ordinateur (ITAO).

L'IAO offre deux principaux avantages par rapport aux autres méthodes de collecte. D'abord, la technique est étayée d'un système de gestion des cas et d'une fonctionnalité de transmission de données. Le système de gestion des cas enregistre automatiquement de l'information de gestion importante sur chaque tentative effectuée dans un cas particulier et produit des rapports aux fins de la gestion du processus de collecte. L'IAO comprend également un ordonnanceur automatique d'appels, c'est-à-dire un système central qui optimise l'horaire des rappels et le calendrier des rendez-vous à l'appui de la collecte par ITAO.

Le système de gestion des cas achemine les applications de questionnaire et les fichiers d'échantillons du Bureau central de Statistique Canada aux bureaux régionaux de collecte dans le cas de l'ITAO et des bureaux régionaux aux ordinateurs portables des intervieweurs dans celui de l'IPAO. Les données destinées au Bureau central sont acheminées en sens inverse. Par souci de confidentialité, les données sont chiffrées avant la transmission. Elles sont ensuite déchiffrées une fois sauvegardées sur un ordinateur sécurisé distinct, sans accès à distance.

Deuxièmement, grâce à l'IAO une interview personnalisée peut être conçue à l'intention de chaque répondant en fonction de ses caractéristiques particulières et des réponses d'enquête. Notamment :

- l'application saute automatiquement les questions qui ne s'appliquent pas au répondant;
- des règles de vérification sont appliquées automatiquement pour repérer les réponses incohérentes ou non incluses dans la fourchette de valeurs permises, et des messages-guides apparaissent à l'écran en réaction à une inscription non valable. De cette façon, l'intervieweur reçoit une rétroaction immédiate et peut corriger toute incohérence;
- le texte des questions, y compris les périodes de référence et les pronoms, est personnalisé automatiquement d'après des facteurs comme l'âge et le sexe du répondant, la date de l'interview et les réponses aux questions précédentes.

6.2 *Développement des applications de l'ESCC*

L'ESCC utilise deux applications d'IAO distinctes pour la collecte de données, l'une pour les interviews téléphoniques (ITAO), l'autre pour les interviews sur place (IPAO). Cette façon de faire permet d'adapter la fonctionnalité de chaque application au type d'interview menée. Chaque application comporte les composantes Entrée, C2 (contenu sur la santé) et Sortie.

Les composantes Entrée et Sortie comprennent des séries standard de questions auxquelles l'intervieweur a pu se référer pour prendre contact avec un répondant, recueillir de l'information importante sur l'échantillon, choisir les répondants et évaluer l'état des cas. La composante C2 consiste en les modules sur la santé et représente la plus grande partie des applications. Il s'agit, notamment, des modules communs posés à tous les répondants et du contenu optionnel, qui variaient d'une région sociosanitaire à l'autre. Chacune des applications a été l'objet de trois étapes de mise à l'essai: les tests modulaires, intégrés et de bout en bout.

Les tests modulaires consistent à mettre à l'essai indépendamment chaque module de contenu afin de vérifier la spécification exacte des instructions « passez à », la logique d'enchaînement et le texte, dans les deux langues officielles. À cette étape, les instructions « passez à » et la logique d'enchaînement entre modules ne sont pas testées, car chaque module est considéré comme un questionnaire autonome. Lorsque les responsables des essais ont terminé la vérification de tous les modules, ces derniers sont regroupés en applications intégrées avec les composantes Entrée et Sortie. À ce moment, les applications intégrées passent à l'étape suivante des essais.

Les tests intégrés portent sur l'ensemble des modules expérimentés, regroupés en applications intégrées avec les composantes Entrée et Sortie. La deuxième étape des essais vise à assurer que des renseignements clés, par exemple l'âge et le sexe, sont transmis de la composante Entrée aux sous-programmes C2 et Sortie. Elle confirme également que les variables qui influent sur les instructions « passez à » et la logique d'enchaînement sont transmises correctement de module en module à l'intérieur de la composante C2. Étant donné que, à ce moment, le fonctionnement des applications est essentiellement identique à ce qu'il sera sur le terrain, tous les scénarios possibles auxquels feront face les intervieweurs sont simulés afin d'en assurer la fonctionnalité rigoureuse. Les scénarios servent à tester divers aspects des composantes Entrée et Sortie, y compris la prise de contact, la collecte d'information sur le contact, la question de savoir si un cas répond à la fourchette des valeurs acceptables, le listage de ménages, la prise de rendez-vous et la sélection de répondants. Les tests servent également à confirmer que, au cours d'une interview, les modules de contenu optionnel choisis pour une région sociosanitaire donnée sont activés.

Les essais de bout en bout situent les applications entièrement intégrées dans un environnement de collecte simulé. Les applications sont chargées dans des ordinateurs connectés à un serveur d'essai. Ensuite, des données sont recueillies, transmises et extraites en temps réel, comme ce serait le cas sur le terrain. Cette dernière étape des essais permet d'expérimenter tous les aspects techniques de la saisie, de la transmission et de l'extraction des données pour chacune des applications de l'ESCC. Il s'agit, par ailleurs, de la dernière occasion de déceler des erreurs dans les composantes Entrée, C2 et Sortie.

6.3 Formation des intervieweurs

Les gestionnaires de projet, les intervieweurs principaux et les intervieweurs des bureaux régionaux responsables de la collecte des données de l'ESCC ont reçu une trousse de formation pour de l'auto-apprentissage avant le début de la collecte. Ces trousse ont été préparées par l'équipe projet de l'ESCC en vue d'être utilisées par les intervieweurs expérimentés de l'ESCC afin de leur procurer une mise à niveau de leurs connaissances. Les gestionnaires de projet et les intervieweurs principaux ont également mené, au besoin, des séances de formation pour les nouveaux intervieweurs. Enfin, des séances de formation mettant l'emphase sur des sujets spécifiques reliés à la collecte des données de l'ESCC se sont déroulées sur une base mensuelle.

L'objet des séances de formation était de familiariser les intervieweurs avec les applications de l'enquête, de les familiariser avec le contenu et de les introduire aux procédures d'entrevue spécifiques à l'ESCC 2009. La formation était centrée sur :

- les buts et objectifs de l'enquête, incluant une partie centrée sur le remaniement de l'enquête;
- les techniques d'enquête;
- les fonctionnalités des applications;
- le contenu des questionnaires, qui a été l'objet d'exercices avec emphase sur les changements significatifs apportés au contenu;
- les techniques à appliquer par l'intervieweur pour mener à bien l'interview, soit des exercices complets dont l'objet était de réduire au minimum la non-réponse;
- la simulation d'interviews difficiles et de situations de non-réponse;
- la gestion de l'enquête;
- les procédures de transmission.

La formation visait en priorité à réduire au minimum les cas de non-réponse. À cette fin, les intervieweurs ont participé à des exercices qui consistaient à persuader des répondants réticents de participer à l'enquête. En outre, les intervieweurs principaux responsables de la conversion des cas de refus à chaque bureau régional de collecte ont participé à une série d'ateliers sur la façon d'éviter les refus.

6.4 L'interview

Des unités d'échantillonnage sélectionnées à partir de la liste de numéros de téléphone et de la base CA ont répondu aux questions posées, à partir de centres d'appel centralisés, par des intervieweurs selon la méthode de l'ITAO. Un intervieweur principal affecté au même centre d'appels assurait la surveillance des intervieweurs. Des intervieweurs sur place décentralisés ont interviewé, au moyen de la méthode de l'IPAO, des unités d'échantillonnage sélectionnées dans la base aréolaire. Bien que, dans certaines situations, les intervieweurs sur place aient été autorisés à mener tout ou partie d'une interview par téléphone, les trois quarts (74,1 %) ont été effectuées exclusivement sur place. Les intervieweurs sur place ont effectué leur travail en autonomie à la maison, au moyen d'ordinateurs portables, et ils étaient surveillés à distance par des intervieweurs principaux. La variable SAM_TYP apparaissant dans les fichiers de microdonnées signifie qu'un cas a été choisi

soit dans la base aréolaire (IPAO), soit dans la liste des numéros de téléphone ou la base CA (ITAO).

Dans tous les logements choisis, l'intervieweur demandait à un membre du ménage bien informé de fournir l'information démographique de base sur tous les occupants. Puis, il sélectionnait un membre du ménage pour une interview plus approfondie, appelée interview C2.

Les intervieweurs qui se servaient de la méthode de l'IPAO ont reçu la formation nécessaire pour procéder à une première prise de contact sur place avec chaque ménage échantillonné. Si la première visite se soldait par une non-réponse, un suivi par téléphone était permis. La variable ADM_N09 apparaissant dans les fichiers de microdonnées indique si l'interview a été effectuée sur place, par téléphone ou au moyen d'un ensemble des deux techniques.

Par souci d'assurer la qualité des données recueillies, les intervieweurs avaient reçu instruction de prendre tous les moyens à leur disposition pour mener en privé l'interview avec le répondant choisi. Là où la situation était inévitable, le répondant a été interviewé en présence d'une autre personne. Dans les fichiers de microdonnées, des indicateurs signalent si une personne autre que le répondant était présente à l'interview (ADM_N10) et si, de l'avis de l'intervieweur, la présence de l'autre personne a influencé les réponses du répondant (ADM_N11).

De nombreuses techniques, y compris les suivantes, ont été mises en œuvre afin de parvenir à un taux de réponse optimal.

a) Lettres d'introduction

Avant le début de chaque période de collecte, les ménages échantillonnés ont reçu des lettres d'introduction qui expliquaient l'objet de l'enquête. Elles énonçaient, notamment, l'importance de l'enquête et offraient des exemples de l'utilisation prévue des données tirées de l'ESCC.

b) Prise de contact

Les intervieweurs ont reçu instruction de mettre en œuvre tous les moyens raisonnables pour obtenir des interviews. Lorsque l'appel (ou la visite) de l'intervieweur était prévu à un moment peu commode, il fixait le moment d'un rappel qui convenait au répondant. Si l'intervieweur ne parvenait pas à prendre rendez-vous par téléphone, il devait effectuer une visite de suivi sur place. S'il n'y avait personne à la maison lors de la première visite, l'intervieweur laissait à la porte une brochure qui expliquait l'enquête et annonçait l'intention de l'intervieweur de prendre contact ultérieurement. De nombreux rappels ont été effectués, à divers moments et différents jours.

c) Conversion des cas de refus

Si une personne refusait d'abord de participer à l'enquête, le bureau régional de Statistique Canada lui faisait parvenir une lettre qui soulignait l'importance de l'enquête et de la collaboration du ménage. Ensuite, un intervieweur principal, un surveillant de projet ou un autre intervieweur rappelait le répondant (ou lui rendait visite) pour faire valoir l'importance de sa participation.

d) Obstacles linguistiques

Pour parer aux problèmes de langue susceptible de nuire aux interviews, tous les bureaux régionaux de Statistique Canada ont embauché des intervieweurs qui parlaient un grand nombre de langues. Au besoin, les cas étaient transférés à un intervieweur capable de remplir le questionnaire dans la langue voulue.

e) Interviews de jeunes

Les intervieweurs ont dû obtenir l'autorisation verbale des parents ou tuteurs pour interviewer des répondants âgés de 12 à 15 ans. Ils ont fait appel à plusieurs procédures pour répondre aux inquiétudes des parents et mener à terme les interviews. Notamment, ils portaient sur eux une fiche intitulée « Note aux parents/tuteurs concernant les interviews de jeunes à l'intention de l'Enquête sur la santé dans les collectivités canadiennes ». La fiche expliquait les raisons pour lesquelles des renseignements étaient recueillis auprès de jeunes, énumérait les thèmes dont traitait l'enquête, demandait l'autorisation de communiquer et de coupler l'information obtenue et expliquait la nécessité de respecter la vie privée et la confidentialité des jeunes.

Si un parent ou tuteur demandait à voir les questions, les intervieweurs avaient pour consigne soit de les leur montrer, soit, si l'interview avait lieu au téléphone, de faire en sorte que le bureau régional leur envoie sur-le-champ un exemplaire du questionnaire.

S'il se révélait impossible d'interviewer en privé le jeune sélectionné, soit sur place, soit par téléphone (sans qu'une autre personne soit à l'écoute), le code de refus était attribué à l'interview. Cependant, dans le cas des interviews selon la méthode de l'IPAO, s'il était impossible d'interviewer en confidence le jeune sélectionné, l'intervieweur pouvait proposer au parent ou tuteur de lui permettre de lire à haute voix les questions, après quoi le jeune pouvait y répondre directement à l'ordinateur.

Lors de toutes les interviews de jeunes, le parent ou tuteur a répondu aux questions sur le revenu et la sécurité alimentaire. Les questions précitées étaient posées à la fin du questionnaire, de telle sorte que le parent ou tuteur pouvait, à ce moment, compléter l'interview.

f) Interviews par procuration

Dans les cas où le répondant sélectionné était, pour des raisons de santé physique ou mentale, incapable de répondre à l'interview, les renseignements à son sujet ont été fournis par un autre membre bien informé du ménage. Cette façon de faire est qualifiée d'interview par procuration. Quoique les interviewés aient été en mesure de donner des réponses exactes à la plupart des questions de l'enquête, les questions plus délicates ou personnelles allaient au-delà des connaissances d'un répondant substitut. Par conséquent, certaines questions posées dans le cadre des interviews par procuration sont demeurées sans réponse. Il fallait donc tout tenter pour réduire au minimum le nombre d'interviews par procuration. La variable ADM_PRX indique si l'interview a été réalisée par procuration ou non.

6.5 Opérations sur le terrain

La plus grande part de l'échantillon de 2009 a été répartie en six périodes de collecte de deux mois chacune qui ne se chevauchent pas. Les bureaux régionaux de collecte ont reçu instruction de passer les quatre premières semaines de chaque période de collecte à interviewer la majorité de l'échantillon, puis de consacrer les quatre semaines suivantes aux interviews restantes et au suivi des cas de non-réponse. À la deuxième semaine de chaque période, des tentatives devaient avoir été effectuées relativement à tous les cas.

Les bureaux de collecte centralisés ont reçu les fichiers d'échantillons environ deux semaines avant le début de chaque période de collecte. Chaque échantillon IPAO comprenait une série de cas fictifs dont devaient s'occuper les intervieweurs principaux afin de confirmer que les procédures de transmission de données fonctionnaient bien tout au long du cycle de collecte. Après réception des échantillons, il incombait aux surveillants de projet de planifier les tâches des intervieweurs chargés des interviews selon la méthode de l'IPAO. Quand la situation s'y prêtait, les tâches étaient limitées à 15 cas par intervieweur.

Le surveillant de projet, l'intervieweur principal et l'équipe de soutien technique du bureau régional étaient chargés de transmettre les cas de chaque bureau responsable des interviews effectuées selon la méthode de l'ITAO au Bureau central. Les transmissions ont eu lieu la nuit, et tous les cas menés à bien ont été acheminés au Bureau central de Statistique Canada. Les interviews menées selon la méthode de l'IPAO ont été transmises chaque jour du domicile de l'intervieweur directement au Bureau central de Statistique Canada par voie d'une ligne de téléphone sécurisée.

Au terme de la collecte de données, un taux de réponse national de 73 % a été atteint. L'annexe E offre le détail des taux de réponse.

6.6 Contrôle de la qualité et gestion de la collecte

Plusieurs méthodes ont servi à assurer la qualité des données et l'optimisation de la collecte des données de 2009. Il s'est agi, entre autres, de mesures internes de vérification du rendement de l'intervieweur et d'une série de rapports de contrôle des diverses cibles de collecte et de la qualité des données.

Le bureau régional validait régulièrement le travail des intervieweurs sur place. Des cas choisis ont été repérés aléatoirement dans les échantillons à chaque période de collecte. Les gestionnaires et les surveillants des bureaux régionaux ont dressé des listes de cas à valider, lesquels ont été confiés à l'équipe de validation, qui communiquait avec les ménages concernés afin de confirmer la tenue d'une interview en règle. Par souci de repérer promptement les problèmes, la validation avait normalement lieu au cours des premières semaines d'une période de collecte. Ensuite, les surveillants adressaient régulièrement une rétroaction aux intervieweurs.

Les intervieweurs travaillant par téléphone ont également été l'objet d'une validation aléatoire. En l'espèce, les intervieweurs principaux des bureaux de collecte responsables des interviews selon la méthode de l'ITAO contrôlaient les interviews afin de vérifier que l'intervieweur appliquait les techniques et les procédures prévues (c'est-à-dire qu'il lisait le libellé des questions tel qu'il figurait dans les applications, qu'il ne posait pas de questions incitatives, et ainsi de suite).

Les responsables ont produit une série de rapports dont l'objet était de contrôler et de gérer efficacement les cibles de collecte et de mettre au jour les problèmes posés par la collecte.

À la fin de chaque période, des rapports cumulatifs ont été produits qui précisaient les taux de réponse, de couplage, de partage et d'interview par procuration, ventilés par échantillon ITAO et IPAO, de même que par région sociosanitaire. Les rapports ont servi à cerner les régions où les niveaux de collecte étaient inférieurs aux cibles, de sorte que les bureaux régionaux puissent y concentrer leurs efforts.

Le Bureau central a effectué des analyses complémentaires au moyen de l'information tirée des applications d'IAO afin de recenser les interviews de durée excessivement courte. Ces dernières ont été signalées au moyen d'indicateurs, supprimées des microdonnées et classées parmi les cas de non-réponse.

7. TRAITEMENT DES DONNÉES

7.1 *Vérification*

La vérification des données a été exécutée en grande partie par l'application d'interview assistée par ordinateur (IAO) durant la collecte des données. Les intervieweurs ne pouvaient pas entrer 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 ».

7.2 *Codage*

On a fourni des catégories de réponses précodées pour toutes les variables appropriées. Les intervieweurs ont reçu une formation durant laquelle ils ont appris à classer les réponses recueillies dans la catégorie appropriée.

Dans les cas où la réponse donnée par le répondant ne pouvait être assignée facilement à une catégorie existante, l'intervieweur pouvait poser plusieurs questions lui permettant d'entrer une réponse en toutes lettres dans la catégorie « Autre – précisez ». Les réponses à toutes ces questions ont été examinées attentivement lors du traitement des données au bureau central. Dans certains cas, on a donné aux réponses en toutes lettres le code d'une catégorie figurant sur la liste si la réponse faisait double emploi. On tiendra compte des réponses « Autre – précisez » fournies pour toutes les questions lors du perfectionnement des catégories de réponses en vue de futurs cycles de l'enquête.

7.3 *Création de variables dérivées*

Pour faciliter l'analyse des données, on a dérivé un certain nombre de variables à partir des éléments disponibles sur le questionnaire de l'ESCC. Le quatrième caractère du nom des variables dérivées est en général un « D », « G » ou un « F ». 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. La *Documentation sur les variables dérivées (VD)* fournit des détails sur la façon de dériver ces variables plus complexes. Pour de plus amples renseignements concernant la nomenclature, veuillez vous référer à la section 12.5.

7.4 Pondération

Le principe de base de l'estimation dans un échantillon aléatoire comme celui de l'ESCC 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 échantillon aléatoire simple 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 à grande diffusion 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 8.

8. PONDÉRATION

Pour que les estimations produites à partir de données d'enquête soient représentatives de la population couverte, et non pas seulement représentatives de l'échantillon comme tel, l'utilisateur doit incorporer les facteurs de pondération, appelés ici les poids d'enquête, dans ses calculs. Un poids d'enquête est attribué à chaque personne incluse dans l'échantillon final, c'est-à-dire dans l'échantillon de personnes ayant répondu à l'enquête. Ce poids correspond au nombre de personnes représentées par le répondant dans l'ensemble de la population de l'enquête.

Tel que décrit dans la section 5, l'ESCC a recours à trois bases de sondage pour la sélection de son échantillon : une base aréolaire de logements agissant comme base principale, puis deux bases formées de numéros de téléphone utilisées pour compléter la base aréolaire. Puisque seulement quelques différences mineures distinguent les deux bases de numéros de téléphone pour la pondération, elles ont été traitées ensemble. On réfère à celles-ci comme faisant partie de la base téléphonique.

Selon les besoins, une seule ou deux bases peuvent être utilisées pour la sélection de l'échantillon dans une région sociosanitaire (RS). Quand on utilise deux bases, la stratégie de pondération traite indépendamment la base aréolaire et la base téléphonique pour dériver les poids au niveau du ménage séparément pour chaque base utilisée. Ces poids-ménages sont ensuite combinés en un seul ensemble de poids-ménages lors d'une étape appelée « *intégration* ». Suite à la transformation des poids-ménages en poids-personnes et à quelques autres ajustements, ce poids intégré devient le poids-personne final.

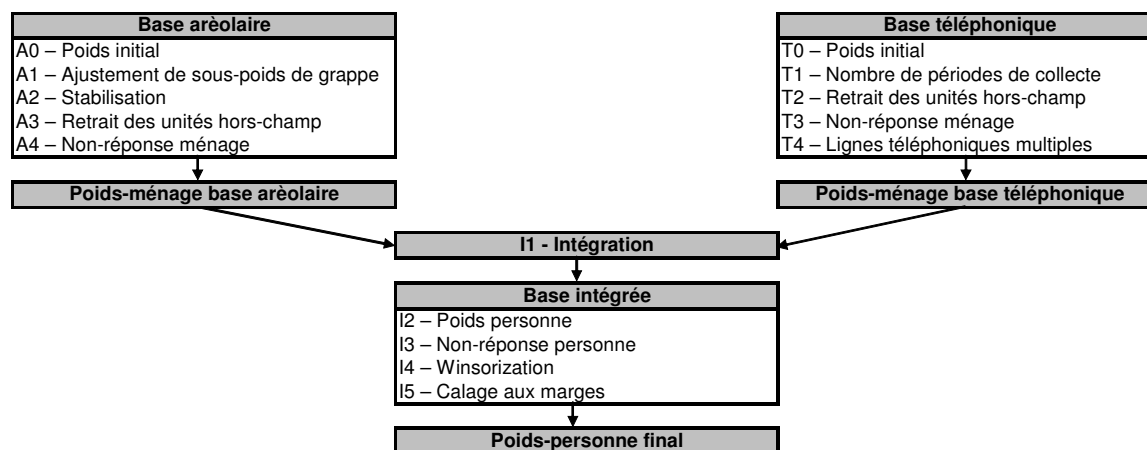
8.1 *Introduction*

Tel que mentionné plus haut, les unités des bases aréolaire et téléphonique sont traitées séparément jusqu'à l'étape d'intégration. Les sections suivantes décrivent la stratégie de pondération pour les provinces. La sous-section 8.2 fournit les détails de la stratégie de pondération pour la base aréolaire, puis la sous-section 8.3, ceux pour la base téléphonique. L'intégration des deux bases est traitée en 8.4. Puis, suivent les deux étapes finales de la pondération, c'est-à-dire l'ajustement pour contrôler la saisonnalité des données puis le calage aux marges, qui sont expliquées dans la sous-section 8.5.

Malgré que les deux bases aient été utilisées pour couvrir les trois territoires, les méthodes d'échantillonnage utilisées ont été légèrement modifiées pour les territoires. Ces modifications affectent substantiellement la pondération pour ces trois régions, et celles-ci sont rapportées dans la sous-section 8.6.

Le diagramme A présente un sommaire des différents ajustements faisant partie de la stratégie de pondération. Un système de numérotation est utilisé pour identifier chaque ajustement apporté au poids et sera utilisé tout au long de la section. Les lettres *A* et *T* sont utilisées comme préfixes pour référer aux ajustements appliqués aux unités des bases Aréolaire et Téléphonique respectivement. Le préfixe *I* est quant à lui utilisé pour identifier l'ajustement d'Intégration et ceux qui suivent.

Diagramme A Sommaire de la stratégie de pondération



8.2 Pondération de l'échantillon provenant de la base aréolaire

A0 – Poids initial

La pondération pour la base aréolaire débute avec un poids fourni par l'Enquête sur la population active (EPA). Ce poids est basé sur le plan de sondage de l'EPA puisque le plan de sondage de base aréolaire de CCHS découle de celui de l'EPA. Le plan de sondage de l'EPA consiste en un échantillonnage de logements dans les grappes sélectionnées parmi les strates de l'EPA. Lors de l'ajustement initial A0, le poids de l'EPA est ajusté afin de tenir compte du fait que l'ESCC sélectionne un échantillon représentatif au niveau des régions sociosanitaires (RS). De plus, l'ESCC sélectionne un nombre de grappes différent de l'EPA et peut répéter l'échantillonnage de logements dans les grappes sélectionnées. Le poids résultant est appelé poids A0. Pour plus de détails sur le mécanisme de sélection, de même qu'une définition plus complète des strates de l'EPA et des grappes, se référer à Statistique Canada (1998)¹⁰.

A1 – Ajustement de sous-poids de grappe

Dans les grappes où un accroissement significatif de la population est observé, une méthode de sous-échantillonnage est utilisée afin de ne pas augmenter indûment la tâche de l'interviewer. La méthode alors appliquée par l'EPA pourra être de sous-échantillonner les logements dans cette grappe, de diviser cette grappe en sous-grappes ou de redéfinir cette grappe comme étant une strate et de créer de nouvelles grappes dans cette strate. Dans tous ces cas, un ajustement de sous-poids de grappe est calculé et appliqué aux poids de CCHS. Cet ajustement est multiplié par le poids A0 pour produire le poids A1. Encore une fois, plus de détails sont disponibles dans la documentation de l'EPA (Statistique Canada (1998)).

¹⁰ Statistique Canada (1998). *Méthodologie de l'enquête sur la population active du Canada*. Statistique Canada numéro 71-526-XPB au catalogue.

A2 – Stabilisation

Dans certaines RS, l'accroissement de l'échantillon, tel que décrit à la section 5, résulte en un échantillon beaucoup plus grand que nécessaire. Une stabilisation a donc été instaurée afin de ramener la taille de l'échantillon au niveau désiré. Le processus de stabilisation consiste à sous-échantillonner des logements aléatoirement à l'intérieur de la RS parmi les logements originalement sélectionnés dans chaque grappe. Un facteur d'ajustement représentant l'effet de la stabilisation est donc calculé afin de corriger la probabilité de sélection. Ce facteur, multiplié par le poids A1, produit le poids A2.

A3 – Retrait des unités hors champ

Parmi tous les logements échantillonnés, une certaine proportion de ceux-ci est, lors de la collecte, identifiée comme étant hors du champ de l'enquête. Des logements détruits ou en construction, des logements vacants, saisonniers ou secondaires, de même que des établissements, sont tous des exemples de cas hors champ pour l'ESCC. Ces logements sont tout simplement retirés de l'échantillon, ne laissant plus que les logements faisant partie du champ de l'enquête. Les logements ou ménages qui restent dans l'échantillon conservent le même poids qu'à l'étape précédente que l'on appelle maintenant poids A3.

A4 – Non-réponse ménage

Lors de la collecte, une certaine proportion des ménages échantillonnés a inévitablement résulté en non-réponse. Ceci survient habituellement lorsque le ménage refuse de participer à l'enquête, fournit des données inutilisables, ou encore, ne peut être rejoint pour réaliser l'interview. Le poids des ménages non-répondants est redistribué aux répondants à l'aide de groupes de réponses homogènes (GRH). Dans le but de créer ces GRH, la méthode du score basée sur une régression logistique est utilisée afin de déterminer la probabilité de réponse, puis ces probabilités sont utilisées afin de diviser l'échantillon en groupes ayant des propriétés de réponse similaire. L'information disponible pour les non-répondants étant limitée, le modèle de régression comprend donc des variables comme la période de collecte et des informations géographiques, en plus de données telles que le nombre d'essais pour contacter le ménage, l'heure/jour des essais et si les essais ont été faits au cours de la semaine ou de la fin de semaine. Depuis 2008, les GRH sont créés à l'intérieur des provinces, pour mieux estimer les totaux provinciaux. Un facteur d'ajustement a donc été calculé à l'intérieur de chaque GRH de la façon suivante:

$$\frac{\text{Somme des poids A3 pour tous les ménages}}{\text{Somme des poids A3 pour tous les ménages répondants}}$$

Le poids A3 des ménages répondants est donc multiplié par ce facteur d'ajustement pour produire le poids A4. Les ménages non-répondants sont éliminés du processus de pondération à partir de ce point.

8.3 Pondération de l'échantillon provenant de la base téléphonique

Tel que mentionné précédemment, la base téléphonique est en fait composée de deux bases : la base de sondage à composition aléatoire (CA), puis une base liste de numéros de téléphone. Une seule de ces deux bases peut être utilisée à l'intérieur d'une RS. La base liste est toujours utilisée comme complément à la base aréolaire tandis que la base CA est toujours utilisée seule pour une RS donnée. Les unités provenant de ces deux bases téléphoniques sont toutefois traitées ensemble et sont donc toutes soumises aux mêmes ajustements.

La géographie utilisée pour sélectionner l'échantillon à partir de la base téléphonique ne répliquait pas parfaitement la géographie des RS, ce qui a forcé certaines unités à être sélectionnées dans une certaine région alors que l'information fournie lors de l'interview les localisait plutôt dans une région avoisinante. Cette particularité a été traitée lors de la pondération en appliquant les trois premiers ajustements (T0, T1 et T2) relativement à la RS assignée lors de la sélection de l'échantillon. Les 2 ajustements restants (T3 et T4) sont appliqués à des RS basé sur l'information recueillie auprès du répondant, pour s'assurer que chaque unité soit associée à la bonne RS.

T0 – Poids initial

Le poids initial est défini comme l'inverse de la probabilité de sélection et il est calculé quelque peu différemment selon que l'échantillon provienne de la base CA ou de la base liste. Dans les deux cas, le poids initial est défini comme étant l'inverse de la probabilité de sélection, mais puisque les méthodes de sélection diffèrent, les probabilités diffèrent aussi. Pour la base CA, la sélection des numéros est faite à l'intérieur de chaque strate CA. Une strate CA représente un agrégat d'indicatifs régionaux et préfixes (IRP: les six premiers chiffres du numéro à 10 chiffres), contenant chacune des banques valides de cent numéros (voir Norris et Paton¹¹ pour plus de détails). Conséquemment, la probabilité de sélection est le ratio entre le nombre d'unités échantillonnées et cent fois le nombre de banques présentes dans la strate CA.

Pour la base liste, les numéros de téléphone sont sélectionnés parmi tous les numéros disponibles dans la liste, et ce indépendamment pour chaque RS. Ainsi, la probabilité de sélection correspond au ratio entre le nombre d'unités échantillonnées et le nombre de numéros de téléphone dans la liste pour la RS. Le ratio est basé sur la version de la base disponible et du nombre d'unités sélectionnées pour une période de collecte donnée. Pour cette raison, la probabilité de sélection peut changer selon la répartition de l'échantillon et les mises à jour de la base. L'inverse de ces probabilités de sélection représente le poids initial T0.

¹¹ Norris, D.A. et Paton, D.G. (1991). L'Enquête sociale générale canadienne: bilan des cinq premières années. *Techniques d'enquête*. 17, 245-260.

T1 – Nombre de périodes de collecte

Contrairement à la base aréolaire, pour laquelle l'échantillon est sélectionné entièrement au début du processus d'échantillonnage, des échantillons sont tirés à chaque deux mois pour les bases téléphoniques. À chacun de ces échantillons bimestriels correspond un poids initial faisant en sorte que chaque échantillon soit représentatif de la RS. Toutefois, pour que l'échantillon total ne représente qu'une seule fois la population, un facteur d'ajustement doit être appliqué pour réduire les poids de chaque échantillon bimestriel. Le facteur d'ajustement appliqué à chaque échantillon bimestriel est égal à l'inverse du nombre d'échantillons combinés (i.e. à l'inverse du nombre de périodes de collecte). À partir de ce moment, l'échantillon de la base liste correspond à la moyenne des échantillons des périodes de collecte combinées. Les poids initiaux sont multipliés par ce facteur d'ajustement pour produire le poids T1.

T2 - Retrait des unités hors champ

Les numéros de téléphone associés à des entreprises, des établissements ou à d'autres logements hors du champ de l'enquête, de même que les numéros hors service sont tous des exemples de cas hors champ pour la base téléphonique. Comme pour la base aréolaire, ces cas sont simplement retirés de l'échantillon, ne laissant ainsi dans l'échantillon que les logements dans le champ de l'enquête. Ces derniers conservent le même poids qu'à l'étape précédente que l'on appelle maintenant poids T2.

T3 – Non-réponse ménage

L'ajustement fait ici pour compenser l'effet de la non-réponse ménage est identique à celui appliqué pour la base aréolaire (ajustement A4). Par contre, les para-données utilisées diffèrent puisque des applications de collecte différentes sont utilisées pour les interviews en personne et par téléphone. C'est donc cette variable qui a été utilisée pour définir les classes d'ajustement. Le facteur d'ajustement calculé à l'intérieur de chaque GRH a été obtenu de la façon suivante:

$$\frac{\textit{Somme des poids T2 pour tous les ménages}}{\textit{Somme des poids T2 pour tous les ménages répondants}}$$

$$\frac{\textit{Somme des poids T2 pour tous les ménages}}{\textit{Somme des poids T2 pour tous les ménages répondants}}$$

Le poids T2 des ménages répondants a donc été multiplié par ce facteur d'ajustement pour produire le poids T3. Les ménages non-répondants sont éliminés à partir de ce point.

T4 - Lignes multiples

Le fait que certains ménages possèdent plus d'une ligne téléphonique résidentielle a un impact sur la pondération: plus le ménage a de lignes, meilleure est sa probabilité d'être sélectionné. Conséquemment, les poids doivent être ajustés pour tenir compte du nombre de lignes résidentielles

que le ménage possède. Le facteur d'ajustement représente l'inverse du nombre de lignes dans le ménage et le poids T4 est obtenu en multipliant ce facteur par le poids T3.

8.4 Intégration des bases aréolaire et téléphonique

Cette étape consiste à intégrer les poids finaux des échantillons aréolaire et téléphonique créés jusqu'à maintenant, en un seul poids en appliquant une méthode d'intégration¹². Le poids des unités qui se trouvent sur la base aréolaire mais pas sur la base téléphonique n'est pas ajusté. Pour toutes les autres unités, un facteur d'ajustement α , compris entre 0 et 1, est appliqué au poids. Le poids des unités de la base aréolaire est multiplié par ce facteur α , alors que le poids des unités de la base téléphonique est multiplié par $1 - \alpha$. Il est à noter que dans les cas où une RS n'est couverte que par une seule base, le facteur d'ajustement est égal à 1. Depuis 2008, un α fixé à 0,4 est utilisé pour les unités qui se trouvent sur les deux bases, dans le but de créer des estimations cohérentes au travers le temps. Le produit du facteur d'ajustement dérivé ici, par le poids de ménage final calculé auparavant (A4 ou T4 dépendant de quelle base provient l'unité), procure le poids intégré I1.

8.5 Les étapes de pondération post-intégration

I2 – Création du poids-personne

Puisque l'unité d'échantillonnage finale pour l'ESCC est la personne, le poids-ménage calculé jusqu'ici doit être converti en un poids-personne. Celui-ci est obtenu en multipliant le poids I1 par l'inverse de la probabilité de sélection de la personne choisie dans le ménage. Nous obtenons ainsi le poids I2. Rappelons que la probabilité de sélection de la personne change en fonction du nombre de personnes dans le ménage et de l'âge des individus (voir section 5.6 pour plus de détails).

I3 – Non-réponse personne

Dans le cadre de l'ESCC, une interview peut être vue comme un processus en deux étapes. Dans un premier temps, l'intervieweur obtient la liste complète des personnes vivant dans le ménage, puis par la suite interviewe la personne sélectionnée dans le ménage. Dans certains cas, les intervieweurs ne réussissent qu'à compléter la première étape, soit parce qu'ils ne peuvent entrer en contact avec la personne sélectionnée, ou encore parce que la personne sélectionnée refuse d'être interviewée. De tels cas sont définis comme étant des non-réponses à l'échelle de la personne, et un facteur d'ajustement doit être appliqué aux poids des personnes répondantes pour compenser cette non-réponse. Tout comme pour la non-réponse à l'échelle du ménage, l'ajustement est appliqué à l'intérieur de groupes de réponses homogènes (GRH). Dans ce processus, la méthode par score est utilisée pour définir une probabilité de réponse définie à partir des caractéristiques disponibles pour les répondants et les non-répondants. Toutes les caractéristiques recueillies lors du listage des membres du ménage, en plus de l'information géographique et des paradosées, sont en fait

¹² Skinner, C.J. et Rao, J.N.K. (1996). Estimation in Dual Frame Surveys with Complex Designs. *Journal of the American Statistical Association*. 91, 433, 349-356.

disponibles pour estimer les probabilités de réponse. Ces probabilités sont utilisées afin de définir les GRH et un facteur d'ajustement est calculé à l'intérieur de chaque GRH de la façon suivante:

$$\frac{\text{Somme des poids I2 pour toutes les personnes sélectionnées}}{\text{Somme des poids I2 pour toutes les personnes sélectionnées répondantes}}$$

Le poids I2 des personnes répondantes est donc multiplié par ce facteur d'ajustement pour produire le poids I3. Les personnes non-répondantes sont éliminées de la pondération à partir de ce point.

I4 – Winsorization

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 RS au point même de devenir aberrants. Ces unités peuvent effectivement représenter une proportion anormalement élevée de leur RS et ainsi influencer fortement les estimations et la variance de ces RS. Afin d'éviter cette situation, le poids des répondants qui contribuent de façon aberrante est ajusté à la baisse selon une méthode appelée « winsorization ».

I5 - Calage aux marges

La dernière étape nécessaire afin d'obtenir le poids final de l'ESCC est le calage aux marges (I5). Le calage aux marges est appliqué en utilisant CALMAR¹³ afin d'assurer que la somme des poids finaux corresponde aux estimations de populations définies à l'échelle des RS, pour chacun des 10 groupes d'âge-sexe d'intérêt, c'est-à-dire les cinq groupes d'âge 12-19, 20-29, 30-44, 45-64, 65+, pour chacun des deux sexes. À partir de 2009, des contrôles additionnels au niveau des sous-régions ont été introduits pour certaines RS. Ces contrôles incluaient les CLSC dans les RS 2403 (Région de la Capitale-Nationale, Québec) et 2415 (Région des Laurentides) et aussi les *District Health Authorities* (DHA) en Nouvelle-Écosse. Dans un même temps, les poids sont ajustés afin de s'assurer que chaque période de collecte (de deux mois) soit également représentée par l'échantillon. Il est à noter que le calage aux marges a été fait en utilisant une géographie la plus à jour possible, qui diffère peut-être de la géographie utilisée lors de l'échantillonnage.

Les estimations de population sont basées sur les plus récents comptes du recensement, de même que sur les comptes de naissance, de décès, d'immigration et d'émigration depuis ce temps. La moyenne des estimations mensuelles pour chacun des croisements RS-âge-sexe par période de collecte a été retenue pour réaliser le calage. Après le calage, l'ajustement de poids I5 est obtenu. Le poids I5 correspond au *poids-personne final de l'ESCC* que l'on retrouve dans le fichier de données portant le nom de variable WTS_M pour les utilisateurs du fichier maître ou du FMGD.

8.6 Particularités de la pondération pour les trois territoires

¹³ Sautory O. CALMAR 2: A New Version of the CALMAR Calibration Adjustment Program. *Proceedings of Statistics Canada Symposium* (Statistics Canada, Catalogue no. 11-522-XCB), 2003.

Tel que décrit à la section 5, le plan d'échantillonnage utilisé pour les trois territoires est quelque peu différent de celui utilisé dans les provinces. La stratégie de pondération est donc adaptée pour répondre à ces différences. Cette section résume les changements apportés à la stratégie expliquée aux sous-sections 8.1 à 8.5.

D'abord pour la base aréolaire, tel que mentionné à la sous-section 5.4.1, une étape additionnelle de sélection est ajoutée pour les territoires. Chaque territoire est initialement stratifié selon des regroupements de communautés à l'intérieur desquels on a sélectionné aléatoirement une communauté. Les capitales de chaque territoire formaient une strate à elles seules, et sont donc toutes trois sélectionnées automatiquement lors de cette première sélection. Cette particularité n'a eu d'effet que dans le calcul de la probabilité de sélection, et donc dans la valeur du poids initial (A0). Une fois ce poids initial calculé, la même série d'ajustements (A1 à A4) est appliquée aux unités de la base aréolaire. Les classes d'ajustement sont construites, pour les non-réponses ménage et personne, à l'aide du même ensemble de variables disponibles pour les provinces.

Pour ce qui est de la pondération des unités de la base téléphonique, mentionnons tout d'abord que seule la base CA est utilisée, et ce, uniquement à l'intérieur des capitales du Yukon et des Territoires du Nord-Ouest. Tous les ajustements de la base téléphonique sont appliqués afin de calculer un poids pour les unités de la base téléphonique.

Les deux ensembles de poids (aréolaire et téléphonique) sont ensuite intégrés, puis poststratifiés de façon semblable à ce qui est fait pour les provinces, à l'exception de trois détails. D'abord, l'intégration a été appliquée uniquement pour les unités situées dans les capitales du Yukon et des Territoires du Nord-Ouest; les autres communautés étant couvertes uniquement par la base aréolaire. Un autre détail à noter pour le Nunavut est que les comptes de la population utilisés pour le calage aux marges représentent les 10 plus grandes communautés (70% de la population) seulement étant donné le sous-dénombrement de la base aréolaire, telle que décrit à la section 5.4.1. Finalement, en commençant avec les produits de diffusion des périodes de référence de 2008 et 2007-2008, des contrôles ont été mis en place pour s'assurer que la proportion des autochtones et la proportion d'individus vivant dans les régions capitales soient contrôlées dans les Territoires du Nord-Ouest et du Yukon. Un contrôle similaire basé sur le statut Inuit a été introduit pour le Nunavut, et à partir de 2009, un contrôle pour la proportion d'individus vivant dans la région capitale de Nunavut est inclus aussi. Ces contrôles assurent que la proportion des estimations représentées par ces groupes est cohérente avec les proportions indiquées par le Recensement de 2006.

8.7 La création d'un poids pour le fichier partagé

En plus du fichier maître et du FMGD qui contiennent tous les répondants de l'ESCC, un fichier partagé est créé qui contient seulement une portion (> 90%) des répondants originaux de l'ESCC. Les individus sur le fichier partagé ont consenti à partager leurs données avec certains partenaires. Pour compenser la perte des répondants sur le fichier, les poids des « partageurs » doivent être ajustés par le facteur :

$$\frac{\text{Somme des poids I5 pour tous les répondants}}{\text{Somme des poids I5 pour les partageurs}}$$

Pareillement à ce qui est fait lors des ajustements pour la non-réponse, ce facteur est créé à l'intérieur de groupes homogènes, en regroupant ici les individus ayant des propensions à partager semblables. Le poids final après cet ajustement est WTS_S.

9. QUALITÉ DES DONNÉES

9.1 *Taux de réponse*

Au total, 84 261 des unités sélectionnées à l'ESCC 2009 faisaient partie du champ de l'enquête¹⁴. Parmi ces unités, 68 526 ont accepté de participer à l'enquête ce qui résulte en un taux de réponse à l'échelle du ménage de 81,3%. Parmi ces ménages répondants, 68 526 personnes ont été sélectionnées (une personne par ménage) pour participer à l'enquête parmi lesquelles 61 679 ont accepté de le faire ce qui résulte en un taux de réponse à l'échelle de la personne de 90,0 %. À l'échelle canadienne, un taux de réponse combiné de 73,2 % a donc été observé pour l'ESCC 2009. Le tableau 9.1 donne les taux de réponse combinés ainsi que l'information pertinente au calcul de ceux-ci pour chaque région sociosanitaire ou regroupement de régions sociosanitaires. Le tableau 9.2 donne la même information mais au niveau des réseaux locaux d'intégration des services de santé (RLISS). Le tableau 9.3 donne le taux de réponse pour le sous-échantillon de l'Enquête sur l'accès aux services de santé (EASS) par province.

Tableau 9.1 : Taux de réponse par région sociosanitaire et par base de sondage de l'ESCC 2009

(voir Annexe E)

Tableau 9.2 : Taux de réponse par réseau local d'intégration des services de santé (RLISS) et par base de sondage de l'ESCC 2009 en Ontario

(voir Annexe E)

Tableau 9.3 : Taux de réponse par province et par base de sondage pour le sous-échantillon de l'Enquête sur l'accès aux services de soins de santé (EASS) de l'ESCC 2009

(voir Annexe E)

On décrit 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.

¹⁴ Parmi les unités sélectionnées au départ, certaines ne font pas partie du champ de l'enquête. Ce sont par exemple des logements vacants ou démolis, des logements non résidentiels ou encore des numéros de téléphone invalides tels les numéros sans service ou non résidentiels. Ces unités sont identifiées pendant la collecte, autrement elles auraient été exclues lors de la sélection. Ces unités ne sont pas considérées dans le calcul des taux de réponse.

Taux de réponse à l'échelle du ménage

Taux de réponse à l'échelle de la personne

Taux de réponse combiné = HHRR x PPRR

Voici maintenant un exemple de calcul du taux de réponse combiné pour le Canada en utilisant l'information fournie dans le tableau 9.1.

$$\begin{aligned}
 \text{Taux de réponse combiné} &= 0,813 \times 0,900 \\
 &= 0,732 \\
 &= \mathbf{73,2 \%}
 \end{aligned}$$

9.2 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.

9.2.1 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 la non-réponse partielle (le fait de ne pas répondre à une ou plusieurs questions) à la non-réponse totale. Dans l'ESCC 2009, il y a peu 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. Dans certaines régions sociosanitaires, l'échantillon initial a été augmenté au cours de l'année afin d'atteindre les tailles ciblées pour 2009. Finalement, les cas de non-réponse totale ont été considérés lors de la pondération en corrigeant les poids des personnes qui ont répondu à l'enquête afin de compenser pour ceux 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.

9.2.2 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 12 ans et plus sont des fumeurs réguliers et que cette estimation comporte un écart type de 0,003. On calcule alors le CV de cette estimation de la façon suivante :

$$(0,003/0,25) \times 100 \% = 1,20 \%$$

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'ESCC 2009 de faire de même. Pour plus d'information sur le calcul des CV, voir la section 11. Pour consulter les lignes directrices sur la façon d'interpréter les résultats du CV, se référer au tableau à la fin de la sous-

10. 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 autre 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.

10.1 Lignes directrices pour l'arrondissement

Afin que les estimations calculées d'après ces fichiers de microdonnées (Maitre, Partager, ou FMGD) 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.

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

Le plan d'échantillonnage utilisé pour cette enquête 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 du fichier 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 au champ du poids, certains progiciels ne permettent pas d'obtenir des estimations qui coïncident exactement avec celles de Statistique Canada.

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

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

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 rentrent dans 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 :

Actuellement, est-ce que ... fume(z) des cigarettes tous les jours, à l'occasion ou jamais?
(SMK_202)

___ Tous les jours

___ À l'occasion

___ Jamais

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 :

Actuellement, combien de cigarettes est-ce que ... fume(z) chaque jour? (SMK_204)

|_| Nombre de cigarettes

10.2.2 Totalisation d'estimations de type nominal

On peut obtenir, à partir du fichier 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.

10.2.3 Totalisation d'estimations quantitatives

Pour obtenir l'estimation d'une somme ou d'une moyenne pour une variable quantitative, on procède aux étapes suivantes (seule l'étape a) est nécessaire pour obtenir l'estimation pour une somme) :

- a) multiplier la valeur de la variable étudiée par le poids finaux puis faire la somme de cette quantité pour tous les enregistrements visés pour 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.

Par exemple, pour estimer le nombre moyen de cigarettes que fument chaque jour les personnes qui fument tous les jours, on calcule d'abord le numérateur (\hat{X}) en sommant le produit entre la valeur de la variable **SMK_204** et le poids **WTS_M**. Ensuite additionnez cette valeur pour les enregistrements pour lesquels la valeur de la variable **SMK_202** est « tous les jours ». On obtient ensuite le dénominateur (\hat{Y}) en additionnant le poids final de tous les enregistrements pour lesquels la valeur de la variable **SMK_202** est « tous les jours ». Le nombre moyen de cigarettes fumées chaque jour par les personnes qui fument tous les jours est finalement obtenu en divisant (\hat{X}) par (\hat{Y}).

10.3 Lignes directrices pour l'analyse statistique

L'ESCC 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), une méthode permet de corriger les résultats obtenus des progiciels courants de façon à ce qu'il soit plus adéquat. Cette méthode consiste à réécherlonner 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.

10.4 Lignes directrices pour la diffusion

Avant de diffuser et/ou de publier des estimations tirées des fichiers de microdonnées, l'utilisateur doit d'abord déterminer le nombre de répondants dans l'échantillon ayant la caractéristique à l'étude (par exemple, le nombre de répondants qui fument lorsqu'on s'intéresse à la proportion de fumeurs pour une population donnée) pour s'assurer qu'il y a assez d'unités pour calculer une estimation de qualité. Pour les utilisateurs de FMGD, si ce nombre est inférieur à 30, l'estimation pondérée ne doit pas être diffusée, quelle que soit la valeur de son coefficient de variation. Pour les utilisateurs des fichiers maître ou de partage, il est recommandé d'avoir au moins 10 observations au numérateur et 20 observations au dénominateur. Pour les estimations pondérées basées sur des

échantillons de 10 ou plus (30 ou plus pour le FMGD), l'utilisateur doit calculer le coefficient de variation de l'estimation arrondie et suivre les lignes directrices qui suivent.

Table 10.1 Lignes directrices relatives à la variabilité d'échantillonnage

Type d'estimation	CV (en %)	Lignes directrices
Acceptable	$0,0 \leq CV \leq 16,5$	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	$CV > 33,3$	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 : « 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. »

11. TABLEAUX DE LA VARIABILITÉ D'ÉCHANTILLONNAGE APPROXIMATIVE

Afin de permettre aux utilisateurs d'avoir facilement accès à des coefficients de variation (CV) qui s'appliqueraient à une multitude d'estimations de type nominal obtenues à partir de ce fichier de microdonnées à grande diffusion, Statistique Canada a produit un ensemble de tableaux de la variabilité d'échantillonnage approximative. Ces tableaux permettent aux utilisateurs d'obtenir un coefficient de variation approximatif selon la taille de l'estimation calculée à partir des données de l'enquête.

Les CV sont calculés en employant la formule de la variance utilisée pour l'échantillonnage aléatoire simple et en y incorporant un facteur qui reflète la structure en grappes à plusieurs degrés du plan d'échantillonnage. Pour obtenir ce facteur, appelé *effet de plan*, on a d'abord calculé les effets de plan pour une vaste gamme de caractéristiques, puis pour chaque tableau, choisi une valeur conservatrice parmi tous les effets de plan relatifs à ce tableau. Cette valeur choisie a ensuite été utilisée pour générer le tableau qui peut alors s'appliquer à l'ensemble complet des caractéristiques.

Les tableaux de la variabilité d'échantillonnage approximative, de même que les effets de plan, les tailles d'échantillon et les comptes de population qui ont servi à les produire, sont disponibles pour les utilisateurs du fichier de partage et du FMGD dans le document *Tableaux de la variabilité d'échantillonnage approximative*. Tous les coefficients de variation sont approximatifs dans les tableaux de la variabilité d'échantillonnage approximative et ils ne doivent donc pas être considérés comme des valeurs exactes. Les possibilités concernant le calcul d'un coefficient de variation exact sont discutées dans la sous-section 11.7.

Rappel : Tel qu'indiqué dans « Les lignes directrices relatives à la variabilité d'échantillonnage » à la section 10.4, si le nombre d'observations sur lesquelles une estimation est basée est inférieur à 30, l'estimation pondérée ne doit pas être diffusée, quelle que soit la valeur de son coefficient de variation. Les coefficients de variation basés sur des échantillons de petite taille sont trop imprévisibles pour être adéquatement représentés dans les tableaux.

11.1 Comment utiliser les tableaux de CV pour les estimations de type nominal

Les règles suivantes devraient permettre à l'utilisateur de calculer, à partir des tableaux de la variabilité d'échantillonnage, les coefficients de variation approximatifs d'estimations relatives au nombre, à la proportion ou au pourcentage de personnes dans la population observée qui possèdent une caractéristique donnée ainsi que des rapports et des écarts entre ces estimations.

Règle 1 : Estimations du nombre de personnes possédant une caractéristique donnée (agrégats)

Le coefficient de variation dépend uniquement de la taille de l'estimation elle-même. Dans le tableau de coefficients de variation approximatifs correspondant à la région appropriée, il faut repérer l'estimation calculée dans la colonne d'extrême gauche (intitulée « Numérateur du pourcentage ») et suivre les astérisques (s'il y en a) de gauche à droite jusqu'au premier nombre. Puisque toutes les valeurs possibles de l'estimation ne sont pas disponibles, il faut prendre la valeur la plus petite qui s'en rapproche le plus (par exemple, si l'estimation vaut 1 700 et que les deux valeurs disponibles sont 1 000 et 2 000, il faut choisir 1 000). Ce nombre constitue le coefficient de variation approximatif pour l'estimation en question.

Règle 2 : Estimations de proportions ou de pourcentages de personnes possédant une caractéristique donnée

Le coefficient de variation d'une proportion (ou d'un pourcentage) estimée dépend à la fois de l'ordre de grandeur de cette proportion et de l'ordre de grandeur du numérateur utilisé dans le calcul de la proportion. Les proportions estimées sont relativement plus fiables que les estimations correspondantes du numérateur de la proportion lorsque celle-ci est fondée sur un sous-ensemble de la population. Cela est dû au fait que les coefficients de variation des estimations du dernier type sont basés sur le chiffre le plus élevé dans une rangée d'un tableau particulier, tandis que les coefficients de variation des estimations du premier type sont basés sur un chiffre quelconque de cette même rangée (pas nécessairement le plus élevé). (Il convient de noter que dans les tableaux, la valeur des coefficients de variation décroît de gauche à droite sur une même ligne.) Par exemple, la proportion estimative de personnes qui fument tous les jours parmi les fumeurs est plus fiable que le nombre estimatif de personnes qui fument tous les jours.

Lorsque la proportion (ou le pourcentage) est fondée sur la population totale de la région géographique à laquelle le tableau s'applique, le coefficient de variation de la proportion est égal à celui du numérateur de la proportion. Dans ce cas-ci, cela équivaut à appliquer la règle 1.

Lorsque la proportion (ou le pourcentage) est fondée sur un sous-ensemble de la population totale (par exemple, les personnes qui fument), il faut se reporter à la proportion (haut du tableau) et au numérateur de la proportion ou du pourcentage (côté gauche du tableau). Puisque toutes les valeurs possibles de la proportion et du numérateur ne sont pas disponibles, il faut, dans les deux cas, prendre la valeur la plus petite qui s'en rapproche le plus (par exemple, si la proportion est de 23 % et que les deux valeurs disponibles dans la colonne s'en rapprochant le plus sont 20 % et 25 %, il faut choisir 20 %). Le coefficient de variation se trouve à l'intersection de la ligne et de la colonne appropriée.

Règle 3 : Estimations des différences entre des agrégats ou des pourcentages

L'erreur-type d'une différence entre deux estimations est à peu près égale à la racine carrée de la somme des carrés de chaque erreur-type considérée séparément. L'erreur-type d'une différence ($\hat{d} = \hat{X}_2 - \hat{X}_1$) est donc :

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

où \hat{X}_1 représente l'estimation 1, \hat{X}_2 l'estimation 2, et α_1 et α_2 sont les coefficients de variation de \hat{X}_1 et \hat{X}_2 respectivement. Le coefficient de variation de \hat{d} est donné par $\sigma_{\hat{d}} / \hat{d}$. Cette formule donne un résultat exact pour ce qui est de la différence entre des sous-populations indépendantes mais n'est autrement qu'approximative. Cette formule mènera à une surestimation de l'erreur si \hat{X}_1 et \hat{X}_2 sont corrélés positivement et à une sous-estimation de l'erreur si \hat{X}_1 et \hat{X}_2 sont corrélés négativement.

Règle 4 : Estimations de rapports

Si le numérateur est un sous-ensemble du dénominateur, il faut convertir le rapport en pourcentage et appliquer la règle 2. Ce serait le cas, par exemple, si le dénominateur est le nombre de personnes qui fument et le numérateur est le nombre de personnes qui fument tous les jours parmi celles qui fument.

Si le numérateur n'est pas un sous-ensemble du dénominateur (par exemple, le rapport du nombre de personnes qui fument tous les jours ou à l'occasion au nombre de personnes qui ne fument pas du tout), l'erreur-type du rapport entre les estimations est à peu près égal à la racine carrée de la somme des carrés de chaque coefficient de variation pris séparément multipliée par \hat{R} , où \hat{R} est le rapport des estimations ($\hat{R} = \hat{X}_1 / \hat{X}_2$). L'erreur-type d'un rapport est donc :

$$\sigma_{\hat{R}} = \hat{R} \sqrt{\alpha_1^2 + \alpha_2^2}$$

où α_1 et α_2 sont les coefficients de variation de \hat{X}_1 et \hat{X}_2 respectivement.

Le coefficient de variation de \hat{R} est donné par $\sigma_{\hat{R}} / \hat{R} = \sqrt{\alpha_1^2 + \alpha_2^2}$. La formule tend à surestimer l'erreur si \hat{X}_1 et \hat{X}_2 sont corrélés positivement et à sous-estimer l'erreur si \hat{X}_1 et \hat{X}_2 sont corrélés négativement.

Règle 5 : Estimations des différences entre des rapports

Dans ce cas-ci, les règles 3 et 4 sont combinées. On commence par calculer les coefficients de variation des deux rapports au moyen de la règle 4, puis le coefficient de variation de leur différence au moyen de la règle 3.

11.2 Exemples d'utilisation des tableaux de CV pour des estimations de type nominal

Les exemples réels suivants ont pour but d'aider les utilisateurs à appliquer les règles décrites ci-dessus.

Exemple 1 : Estimations du nombre de personnes possédant une caractéristique donnée (agrégats)

Supposons qu'un utilisateur estime à 4 722 617 le nombre de personnes qui fument tous les jours au Canada. Comment l'utilisateur fait-il pour déterminer le coefficient de variation de cette estimation?

- 1) Se reporter au tableau de CV pour le CANADA.
- 2) L'agrégat estimé (4 722 617) ne figure pas dans la colonne de gauche (la colonne « Numérateur du pourcentage »); il faut donc utiliser le nombre le plus petit qui s'en rapproche le plus, soit 4 000 000.
- 3) Le coefficient de variation d'un agrégat estimé (exprimé en pourcentage) est la première entrée sur cette ligne (à part les astérisques), soit 1,7 %.
- 4) Le coefficient de variation approximatif de l'estimation est donc 1,7 %. Par conséquent, selon les « Lignes directrices relatives à la variabilité d'échantillonnage » présentées à la section 10.4, l'estimation selon laquelle 4 722 617 personnes fument tous les jours peut être diffusée sans réserve.

Exemple 2 : Estimations de proportions ou de pourcentages de personnes possédant une caractéristique donnée

Supposons qu'un utilisateur estime à $4\,722\,617/6\,081\,453 = 77,7\%$ le pourcentage de personnes, parmi les fumeurs, qui fument tous les jours au Canada. Comment l'utilisateur fait-il pour déterminer le coefficient de variation de cette estimation?

- 1) Se reporter au tableau de CV pour le CANADA.
- 2) Parce que l'estimation est un pourcentage basé sur un sous-ensemble de la population totale (c.-à-d. les personnes qui fument tous les jours ou à l'occasion), il faut utiliser à la fois le pourcentage (77,7 %) et la partie numérateur du pourcentage (4 722 617) pour déterminer le coefficient de variation.

- 3) Le numérateur (4 722 617) ne figure pas dans la colonne de gauche (la colonne « Numérateur du pourcentage »); il faut donc utiliser le nombre le plus petit qui s'en rapproche le plus, soit 4 000 000. De même, l'estimation du pourcentage ne figure pas parmi les en-têtes de colonnes; il faut donc utiliser le nombre le plus petit qui s'en rapproche le plus, soit 70,0 %.
- 4) Le nombre qui se trouve à l'intersection de la ligne et de la colonne utilisées, soit 1,0 %, est le coefficient de variation (exprimé en pourcentage) à employer.
- 5) Le coefficient de variation de l'estimation est donc 1,0 %. Par conséquent, selon les « Lignes directrices relatives à la variabilité d'échantillonnage » présentées à la section 10.4, l'estimation selon laquelle 77,7 % des gens qui fument le font tous les jours peut être diffusée sans réserve.

Exemple 3 : Estimations des différences entre des agrégats ou des pourcentages

Supposons qu'un utilisateur estime que, parmi les hommes, $2\,535\,367/13\,078\,499 = 19,4\%$ fument tous les jours (estimation 1), alors que chez les femmes, ce pourcentage est estimé à $2\,187\,250/13\,476\,931 = 16,2\%$ (estimation 2). Comment l'utilisateur fait-il pour déterminer le coefficient de variation de la différence entre ces deux estimations?

- 1) À l'aide du tableau de CV pour le CANADA, utilisé de la même façon que dans l'exemple 2, vous établissez à 2,4 % le CV de l'estimation 1 (exprimé en pourcentage) et à 2,4 % le CV de l'estimation 2 (exprimé en pourcentage).
- 2) Selon la règle 3, l'erreur-type pour une différence ($\hat{d} = \hat{X}_2 - \hat{X}_1$) est :

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

où \hat{X}_1 est l'estimation 1, \hat{X}_2 est l'estimation 2, et α_1 et α_2 sont les coefficients de variation de \hat{X}_1 et \hat{X}_2 respectivement. L'erreur-type de la différence $\hat{d} = (0,194 - 0,162) = 0,032$ est donc :

$$\begin{aligned} \sigma_{\hat{d}} &= \sqrt{[(0,194)(0,024)]^2 + [(0,162)(0,024)]^2} \\ &= 0,0061 \end{aligned}$$

- 3) Le coefficient de variation de \hat{d} est donné par $\sigma_{\hat{d}} / \hat{d} = 0,0061/0,032 = 0,190$.
- 4) Le coefficient de variation approximatif de la différence entre les estimations est donc 19,0 % (exprimé en pourcentage). Par conséquent, toujours selon les « Lignes directrices relatives à la variabilité d'échantillonnage » présentées à la section 10.4, cette estimation peut être publiée en y joignant une mise en garde.

Exemple 4 : Estimations de rapports

Supposons qu'un utilisateur estime à 4 722 617 le nombre de personnes qui fument tous les jours et à 1 358 836 le nombre de celles qui fument à l'occasion. L'utilisateur veut comparer ces deux estimations sous la forme d'un rapport. Comment fait-il pour déterminer le coefficient de variation de cette estimation?

- 1) Tout d'abord, cette estimation est une estimation de rapport, où le numérateur de l'estimation ($= \hat{X}_1$) est le nombre de personnes qui fument à l'occasion. Le dénominateur de l'estimation ($= \hat{X}_2$) est le nombre de personnes qui fument tous les jours.
- 2) Se reporter au tableau de CV pour le CANADA.
- 3) Le numérateur de cette estimation de rapport est 1 358 836. Le nombre le plus petit qui se rapproche le plus de ce nombre est 1 000 000. Le coefficient de variation de cette estimation (exprimé en pourcentage) est la première entrée sur cette ligne (à part les astérisques), soit 3,7 %.
- 4) Le dénominateur de cette estimation de rapport 4 722 617. Le nombre le plus petit qui se rapproche le plus de ce nombre est 4 000 000. Le coefficient de variation de cette estimation (exprimé en pourcentage) est la première entrée sur cette ligne (à part les astérisques), soit 1,7 %.
- 5) Le coefficient de variation approximatif de l'estimation du rapport est donc donné par la règle 4,

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^2 + \alpha_2^2},$$

c'est-à-dire,

$$\alpha_{\hat{R}} = \sqrt{(0,037)^2 + (0,017)^2}$$

$$= 0,041$$

où α_1 et α_2 sont les coefficients de variation de \hat{X}_1 et \hat{X}_2 respectivement. Le rapport des personnes qui fument occasionnellement à celles qui fument tous les jours est 1 358 836/4 722 617, soit 0,29:1. Le coefficient de variation de cette estimation est 4,1 % (exprimé en pourcentage); selon les « Lignes directrices relatives à la variabilité d'échantillonnage » présentées à la section 10.4, l'estimation peut donc être diffusée sans réserve.

11.3 Comment utiliser les tableaux de CV pour calculer les limites de confiance

Bien que les coefficients de variation soient largement utilisés, l'intervalle de confiance d'une estimation représente une mesure plus intuitive de l'erreur d'échantillonnage. Un intervalle de confiance est une façon d'énoncer la probabilité que la valeur vraie de la population se situe dans une plage de valeurs données. Par exemple, un intervalle de confiance de 95 % peut être décrit comme suit : si l'échantillonnage de la population se répète à l'infini, chacun des échantillons donnant un nouvel intervalle de confiance pour une estimation, l'intervalle contiendra la valeur vraie de la population dans 95 % des cas.

Une fois déterminée l'erreur-type d'une estimation, on peut calculer des intervalles de confiance pour les estimations en partant de l'hypothèse qu'en procédant à un échantillonnage répété de la population, les diverses estimations obtenues pour une caractéristique de la population sont réparties selon une distribution normale autour de la valeur vraie de la population. Selon cette hypothèse, il y a environ 68 chances sur 100 que l'écart entre une estimation de l'échantillon et la valeur vraie de la population soit inférieur à une erreur-type, environ 95 chances sur 100 que l'écart soit inférieur à deux erreurs-types et environ 99 chances sur 100 que l'écart soit inférieur à trois erreurs-types. On appelle ces différents degrés de confiance des niveaux de confiance.

L'intervalle de confiance d'une estimation \hat{X} est généralement exprimé sous la forme de deux nombres, l'un étant inférieur à l'estimation et l'autre supérieur à celle-ci, sous la forme $(\hat{X} - k, \hat{X} + k)$, où k varie selon le niveau de confiance désiré et l'erreur d'échantillonnage de l'estimation.

On peut calculer directement les intervalles de confiance d'une estimation à partir des tableaux de la variabilité d'échantillonnage approximative, en trouvant d'abord dans le tableau approprié le coefficient de variation de l'estimation \hat{X} , puis en utilisant la formule suivante pour obtenir l'intervalle de confiance CI correspondant :

$$CI_X = [\hat{X} - z \hat{X} \alpha_{\hat{X}}, \hat{X} + z \hat{X} \alpha_{\hat{X}}]$$

où $\alpha_{\hat{X}}$ est le coefficient de variation trouvé pour \hat{X} , et

$z = 1$ si l'on désire un intervalle de confiance de 68 %

$z = 1,6$ si l'on désire un intervalle de confiance de 90 %

$z = 2$ si l'on désire un intervalle de confiance de 95 %

$z = 3$ si l'on désire un intervalle de confiance de 99 %

Nota : Les lignes directrices concernant la diffusion des estimations de la section 10.4 s'appliquent aussi aux intervalles de confiance. Par conséquent, si l'estimation ne peut être diffusée, alors l'intervalle de confiance ne peut l'être lui non plus.

11.4 Exemple d'utilisation de tableaux de CV pour obtenir des limites de confiance

Voici la marche à suivre pour calculer un intervalle de confiance de 95 % pour la proportion estimée de personnes qui fument tous les jours parmi celles qui fument (d'après l'exemple 2 de la sous-section 11.2).

$$\hat{X} = 0,777$$

$$z = 2$$

$$\alpha_{\hat{x}} = 0,01 \text{ est le coefficient de variation de cette estimation selon les tableaux.}$$

$$CI_{\hat{x}} = \{0,777 - (2) (0,777) (0,01), 0,777 + (2) (0,777) (0,01)\}$$

$$CI_{\hat{x}} = \{0,761, 0,793\}$$

11.5 Comment utiliser les tableaux de CV pour effectuer un test Z

On peut aussi utiliser les erreurs-types pour effectuer des tests d'hypothèses, une technique qui permet de faire la distinction entre les paramètres d'une population à l'aide d'estimations basées sur un échantillon. Ces estimations peuvent être des nombres, des moyennes, des pourcentages, des rapports, etc. Les tests peuvent être effectués à divers niveaux de signification; un niveau de signification est la probabilité de conclure que les caractéristiques sont différentes quand, en fait, elles sont identiques.

Supposons que \hat{X}_1 et \hat{X}_2 sont des estimations basées sur un échantillon pour deux caractéristiques voulues. Supposons aussi que l'erreur-type de la différence $\hat{X}_1 - \hat{X}_2$ est $\sigma_{\hat{a}}$. Si $z = (\hat{X}_1 - \hat{X}_2) / \sigma_{\hat{a}}$ est compris entre -2 et 2, alors on ne peut tirer aucune conclusion à propos de la différence entre les caractéristiques au niveau de signification de 5 %. Toutefois, si ce rapport est inférieur à -2 ou supérieur à +2, la différence observée est significative au niveau de 0,05.

11.6 Exemple d'utilisation des tableaux de CV pour effectuer un test Z

Supposons que nous voulons tester, au niveau de signification de 5 %, l'hypothèse selon laquelle il n'y a pas de différence entre la proportion d'hommes qui fument tous les jours et cette même proportion chez les femmes. Dans l'exemple 3 de la sous-section 11.2, nous avons déterminé que l'erreur-type de la différence entre ces deux estimations est égale à 0,0061. Par conséquent,

$$z = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_d} = \frac{0,194 - 0,162}{0,0061} = \frac{0,032}{0,0061} = 5,25$$

Puisque $z = 5,25$ est supérieur à 2, on doit conclure qu'il existe une différence significative entre les deux estimations au niveau de signification de 0,05. À noter que les deux sous-groupes comparés sont considérés comme étant indépendants, ce qui fait en sorte que le résultat du test est valide.

11.7 Variances ou coefficients de variation exacts

Tous les coefficients de variation qui figurent dans les tableaux de la variabilité d'échantillonnage approximative (tableaux de CV) sont effectivement approximatifs, donc, non officiels.

Le calcul de variance ou coefficient de variation exact n'est pas chose évidente puisqu'il n'existe pas de formule mathématique simple pouvant prendre en compte tous les aspects du plan d'échantillonnage et de la pondération de l'ESCC. On doit donc avoir recours à d'autres méthodes pour estimer ces mesures de précisions, telles que des méthodes par rééchantillonnage. Parmi celles-ci, la méthode du bootstrap est celle recommandée pour l'analyse des données de l'ESCC.

Le calcul de coefficients de variation (ou tout autre mesure de précision) fait à l'aide de la méthode du bootstrap nécessite toutefois l'accès à de l'information considérée confidentielle qui n'est évidemment pas disponible dans le fichier de microdonnées à grande diffusion. Le calcul doit donc se faire à l'aide du fichier maître. L'accès au fichier maître est discuté à la section 12.3.

Pour le calcul de coefficients de variation, il est conseillé d'utiliser la méthode du bootstrap. Un programme macro, appelé le « Bootvar », a été développé pour faciliter le calcul à l'aide de la méthode bootstrap. Le programme Bootvar est offert en formats SAS et SPSS, et est constitué de macros qui calculent les variances de totaux, ratios, différences entre ratios, et pour des régressions linéaires et logistiques.

Les raisons pour lesquelles un utilisateur pourrait souhaiter connaître la précision exacte de ses estimations sont diverses. En voici quelques-unes.

Premièrement, si un utilisateur désire obtenir des estimations à un niveau géographique autre que ceux présentés dans les tableaux (par exemple, au niveau urbain ou rural), l'utilisation des tableaux de CV publiés ne convient pas parfaitement. Néanmoins, on peut obtenir les coefficients de variation de ce type d'estimations en appliquant la méthode d'estimation par domaine, au moyen du programme de calcul de la variance exacte (le « Bootvar »).

Deuxièmement, si un utilisateur demande des analyses plus complexes, telles que des estimations de paramètres de modèles de régression linéaire ou logistique, les tableaux de CV ne pourront pas fournir les coefficients de variation pour ceux-ci. Certains progiciels statistiques courants permettent d'incorporer les poids d'échantillonnage aux analyses, mais, souvent, les variances produites ne tiennent pas bien compte de la stratification et de la mise en grappe de l'échantillon, contrairement à celles obtenues grâce au programme de calcul de la variance exacte.

Troisièmement, dans le cas de l'estimation de variables quantitatives, il est nécessaire d'utiliser des tableaux distincts pour déterminer l'erreur d'échantillonnage. Or, la plupart des variables de l'ESCC étant de type nominal, de tels tableaux n'ont pas été produits. Les utilisateurs qui souhaitent connaître les coefficients de variation de variables quantitatives peuvent néanmoins obtenir ces derniers grâce au programme de calcul de la variance réelle. À noter, toutefois, que le coefficient de variation d'un total quantitatif est généralement plus grand que celui de l'estimation de type nominal correspondante (c'est-à-dire, l'estimation du nombre de personnes qui contribuent à l'estimation quantitative). Si l'estimation de type nominal correspondante ne peut être diffusée, il en sera de même pour l'estimation quantitative. Par exemple, le coefficient de variation de l'estimation du nombre total de cigarettes que fument chaque jour les personnes qui fument tous les jours serait supérieur à celui de l'estimation correspondante du nombre de personnes qui fument tous les jours. Par conséquent, si on ne peut diffuser le coefficient de variation de cette dernière estimation, on ne pourra non plus diffuser celui de l'estimation quantitative correspondante.

Enfin, un utilisateur qui peut se servir des tableaux de CV, mais obtient ainsi un coefficient de variation compris dans la fourchette marginale (de 16,6 % à 33,3 %), devrait diffuser les estimations associées en y joignant une mise en garde aux utilisateurs quant à la variabilité d'échantillonnage élevée liée aux estimations. Dans ce cas, il serait bon de recalculer le coefficient de corrélation à l'aide du programme de variance exacte pour vérifier si ces estimations peuvent être diffusées sans mise en garde. Cette situation tient au fait que l'estimation des coefficients de variation grâce aux tableaux de la variabilité d'échantillonnage approximative est basée sur une vaste gamme de variables et est donc jugée grossière, tandis que le programme de calcul de la variance réelle produit le coefficient de variation précis associé à la variable en question.

11.8 Seuils pour la diffusion des estimations relatives à l'ESCC

Le document *Tableaux de la variabilité d'échantillonnage approximative*, qui est disponible pour les utilisateurs du fichier de partage et du FMGD, présente les tableaux indiquant les seuils de diffusion des totaux selon les estimations pour le Canada, les provinces, les régions sociosanitaires et les CLSC, ainsi que pour les différents groupes d'âge (pour le Canada seulement). Les estimations inférieures à la valeur indiquée dans la colonne « Marginal » ne peuvent en aucun cas être diffusées.

12 FICHIERS DE MICRODONNÉES : DESCRIPTION, ACCÈS ET UTILISATION

L'ESCC produit trois types de fichiers de microdonnées : les fichiers maîtres, les fichiers de partage et les fichiers de microdonnées à grande diffusion (FMGD). Le tableau 12.1 contient la liste des fichiers disponibles pour les données de 2009.

12.1 Fichiers maîtres

Les fichiers maîtres contiennent toutes les variables et tous les enregistrements de l'enquête collectés au cours d'une période de collecte. Ces fichiers sont accessibles à Statistique Canada pour usage interne, dans les Centres de données de recherche (CDR) de Statistique Canada et peuvent aussi faire l'objet de demandes de totalisations personnalisées.

12.1.1 Centre de données de recherche

Le Programme des CDR permet aux chercheurs d'utiliser les données d'enquête contenues dans les fichiers maîtres dans un environnement sécuritaire situé dans plusieurs universités à travers le Canada. Les chercheurs doivent soumettre des propositions de recherche qui une fois acceptées leur donneront accès aux CDR. Pour plus de renseignements, consultez la page web suivante : http://www.statcan.ca/francais/rdc/index_f.htm

12.1.2 Totalisations personnalisées

Une autre méthode d'accès aux fichiers maîtres consiste à offrir à tous les utilisateurs de faire appel au personnel du Service à la clientèle de la Division de la statistique de la santé pour produire des totalisations personnalisées. Ce service est offert moyennant le recouvrement des coûts. Il permet aux utilisateurs qui ne savent pas se servir de logiciels de totalisation d'obtenir des résultats personnalisés. Les résultats sont filtrés pour s'assurer qu'ils sont conformes aux normes de confidentialité et de fiabilité avant d'être diffusés. Pour plus de renseignements, communiquez avec le Service à la clientèle (613) 951-1746 ou par courriel à hd-ds@statcan.gc.ca.

12.1.3. Télé-accès

En dernier lieu, le service de télé-accès aux fichiers maîtres de l'enquête est un moyen d'accéder à ces données s'il est impossible de passer par un Centre de recherche en données. On peut fournir à l'acheteur d'un produit de microdonnées un fichier maître de données synthétique ou fichier « fictif » et le cliché d'enregistrement correspondant. Grâce à ces outils, le chercheur peut mettre au point son propre ensemble de programmes analytiques. Il ne lui reste plus qu'à envoyer le code pour les totalisations personnalisées par courrier électronique à cchs-escc@statcan.gc.ca. Le code est transmis au réseau interne protégé de Statistique Canada et traité en regard du fichier maître approprié de données de l'ESCC. Les estimations générées seront communiquées à l'utilisateur, sujet aux directives sur l'analyse et la communication des données tel qu'exposé dans les grandes lignes à la section 10 de ce document. Les résultats sont filtrés pour vérifier qu'ils sont conformes aux normes de confidentialité et de fiabilité, puis, les données de sortie sont renvoyées au client. Ce service est gratuit.

12.2 Fichiers de partage

Les fichiers de partage contiennent toutes les variables et tous les enregistrements des répondants de l'ESCC qui ont accepté de partager leurs données avec les partenaires de partage de Statistique Canada, soit les ministères de la santé des provinces et territoires, Santé Canada et l'Agence de santé de la fonction publique. Statistique Canada demande également aux répondants résidant au Québec leur permission de partager leurs données avec l'Institut de la statistique du Québec. Statistique Canada ne fournit le fichier de partage qu'à ces organisations. Les identificateurs personnels sont retirés des fichiers de partage pour préserver la confidentialité des répondants. Les utilisateurs de ces fichiers doivent au préalable avoir porté serment qu'ils ne divulgueront en aucun temps toute information susceptible d'identifier un répondant à l'enquête.

12.3 Fichiers de microdonnées à grande diffusion

Les fichiers de microdonnées à grande diffusion (FMGD) sont élaborés à partir des fichiers maîtres suivant une technique qui vise à concilier l'impératif d'assurer la confidentialité des répondants et la nécessité de produire des données d'utilité maximale à l'échelle de la région sociosanitaire. Les FMGD doivent répondre à des normes sévères de sécurité et de confidentialité, conformément à la *Loi sur la statistique* avant qu'ils ne soient diffusés pour l'accès public. Pour s'assurer du respect de ces normes, chaque FMGD est soumis à un processus officiel d'examen et d'approbation par un comité formé de haut gestionnaire de Statistique Canada.

Les variables les plus susceptibles de permettre l'identification d'une personne sont supprimées du fichier ou agrégées en catégories moins détaillées.

Le FMGD contient les données des questions posées sur une période de deux ans. À moins d'exception, ces questions sont habituellement celles comprises dans le contenu commun annuel et dans le contenu commun deux ans, ainsi que dans le contenu optionnel, choisi pour deux ans par les provinces et les territoires.

Les FMGD sont accessibles gratuitement dans les établissements d'enseignement post-secondaires font partie de l'Initiative de démocratisation des données. Ils sont aussi disponibles gratuitement sur demande auprès du Service à la clientèle au 613-951-1746 ou à hd-ds@statcan.gc.ca.

Tableau 12.1 Fichiers de données de l'ESCC 2009

Fichiers	Nom des fichiers	Poids d'échantillonnage	Fichiers des poids bootstrap	Variables incluses	Enregistrements inclus
Fichier maître principal	HS.txt	WTS_M	b5.txt	Tous les modules communs annuels et un an.	Enregistrements de tous les répondants
Fichier maître du	HSS1.txt	WTS_S1M	b5_s1.txt	Tous les modules communs, plus	Enregistrements de tous les répondants

sous-Échantillon 1				l'ensemble des variables « Z » pour les modules Accès aux services de santé et Temps d'attentes.	sélectionnés pour le sous-échantillon.
Fichier de partage	HS.txt	WTS_S	b5.txt	Tous les modules communs et optionnels.	Enregistrements de tous les répondants qui ont accepté de partager leurs données
Fichier de partage du sous-échantillon 1	HSS1.txt	WTS_S1S	b5_s1.txt	Tous les modules communs, plus l'ensemble des variables « Z » pour les modules Accès aux services de santé et Temps d'attentes.	Enregistrements de tous les répondants sélectionnés pour le sous-échantillon qui ont accepté de partager leurs données

12.4 Utilisation des fichiers de l'ESCC : fichier annuel ou fichier deux ans ?

Depuis les diffusions des données de 2008 et de 2007-2008, les utilisateurs qui ont accès aux fichiers de partage ou aux fichiers maîtres ont le choix d'utiliser les fichiers d'un an ou de deux ans. Les décisions concernant la période à utiliser pour une analyse donnée devraient être guidées par le niveau de détail et de qualité requis. Dans le cas d'un fichier d'un an, les estimations ne sont pas toujours disponibles en raison de la qualité liée à la taille limitée de l'échantillon.

Dans le cadre de l'ESCC, avant d'interpréter et d'utiliser une estimation, il est recommandé de s'assurer que cette estimation rencontre les règles suivantes :

- coefficient de variation 33.3 % ou moins ;
- minimum de 10 répondants ayant la caractéristique dans ce domaine (numérateur);
- total minimum de 20 répondants pour le domaine d'intérêt (dénominateur).

Cela ne sera pas possible dans le cas des caractéristiques rares et des domaines détaillés pour les fichiers d'un an. Les utilisateurs devront plutôt se fier aux fichiers de deux ans ou aux fichiers pluriannuels.

Lorsque l'utilisation d'un fichier d'un an ou de deux ans est possible, l'utilisateur devrait envisager un compromis entre l'exactitude et l'actualité. S'il est important de rendre compte des caractéristiques courantes d'une population le plus précisément possible, les fichiers d'un an sont préférables. Toutefois, du fait de l'augmentation de la taille de l'échantillon, des estimations et des analyses plus détaillées peuvent être effectuées avec les fichiers deux ans.

12.5 Utilisation de la variable de pondération

La variable de pondération **WTS_M** représente le poids d'échantillonnage pour les fichiers principaux de l'enquête. Pour un répondant donné, ce poids d'échantillonnage peut être interprété comme étant le nombre de personnes que le répondant représente dans la population canadienne. Ce poids doit être utilisé en tout temps dans les calculs d'estimations statistiques, afin de permettre l'inférence à l'échelle de la population. La production de résultats non pondérés n'est pas recommandée. La répartition de l'échantillon, de même que les détails du plan de sondage, peuvent entraîner des résultats biaisés qui ne représentent pas correctement la population. Pour une description plus détaillée du calcul de ce poids, consulter la section 8 sur la pondération. La variable de pondération **WTS_M** doit être utilisée pour des analyses régionales.

Le module Sécurité alimentaire inclut dans les fichiers de certaines périodes référence, mesure des concepts qui s'appliquent non seulement à la situation du répondant, mais à celle de l'ensemble de son ménage. L'analyse des variables peut, selon le niveau d'analyse, nécessiter le recours à un poids calculé de manière à représenter le nombre de ménages au Canada, plutôt que le nombre de personnes. Cette variable de pondération **WTS_HH** se trouve dans un fichier distinct

(HS_HHWT.txt). Elle peut être utilisée en remplacement de la variable **WTS_M** pour des analyses au niveau des ménages à l'échelle nationale et provinciale.

12.6 Convention appliquée pour nommer les variables à partir de 2007

Les conventions appliquées pour nommer des variables permettent aux utilisateurs des données de repérer et d'utiliser facilement celles-ci en fonction du module et du type de variable. La convention appliquée pour nommer les variables de l'ESCC respectent deux exigences: limiter les noms des variables à huit caractères au plus pour qu'il soit facile de les utiliser avec les logiciels d'analyse et permettre de repérer facilement les variables conceptuellement identiques d'une période de collecte à l'autre de l'enquête. Les questions auxquelles on a apporté des changements entre deux périodes de collecte, qui modifient le concept mesuré par la question, sont entièrement renommées pour éviter toute confusion dans l'analyse.

La convention appliquée pour nommer les variables de l'ESCC a été modifiée à partir de 2007. Ainsi, la lettre correspondant à l'édition de l'enquête (par exemple, A = 2000 (cycle 1.1), C = 2003 (cycle 2.1) et E = 2005 (3.1) n'est plus utilisée dans les noms de variables. Une nouvelle variable (REFPER, format = AAAAMM-AAAAMM) a été ajoutée aux fichiers de microdonnées afin d'identifier le début et la fin de la période de référence au cours de laquelle les données ont été recueillies. Celle-ci s'avérera utile en particulier pour les utilisateurs désireux d'utiliser les données de plusieurs périodes de collecte à la fois. Ainsi, les noms des variables correspondant à des modules ou à des questions identiques entre une année de collecte et une autre (par ex., 2007 et 2008) seront les mêmes.

La convention appliquée pour nommer les variables à compter de l'ESCC 2007 utilise jusqu'à huit caractères. Les noms de variables sont structurés de la manière suivante :

Positions 1 à 3 : Nom du module/de la section du questionnaire
Position 4 : Type de variable (sous-tiret, C, D, F ou G)
Positions 5 à 8 : Numéro de la question et option de réponse (s'il s'agit d'une question à réponse multiple)

L'exemple 1 présente la structure du nom de la variable correspondant à la question 202 du module Usage du tabac, c'est-à-dire SMK_202 :

Positions 1 à 3 : SMK Module sur l'usage du tabac
Position 4 : _ (sous-tiret = données recueillies)
Position 5 à 8 : 202 numéro de la question

L'exemple 2 présente la structure du nom de la variable correspondant à la question 2 du module Utilisation des soins de santé (HCU_02A) qui constitue une question à réponse multiple :

Positions 1 à 3 : HCU Module sur l'utilisation des soins de santé
Position 4 : _ (sous-tiret = données recueillies)
Position 5 à 8 : 02AA numéro de la question correspondant et option de réponse

Dans les positions 1 à 3, on retrouve l'acronyme de chacun des modules. Ces acronymes apparaissent à côté des noms de modules qui sont tous présentés dans la figure de l'annexe A.

La position 4 désigne le type de variable selon qu'il s'agit d'une variable collectée directement à partir d'une question du questionnaire (« _ »), d'une variable codée (« C »), dérivée (« D »), groupée (« G ») ou d'une variable indicatrice (« F »)¹⁵.

En général, les quatre dernières positions (5 à 8) 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 ou trois chiffres. Par exemple, la question Q01A du questionnaire devient simplement 01A et la question Q15, simplement 15.

¹⁵ Du mot anglais *flag*.

Tableau 12.2 Désignation des codes utilisés à la 4e position du nom des variables de l'ESCC

–	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 collectées 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., groupes d'âge)

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.

12.7 Convention appliquée pour nommer les variables avant 2007

Tel que mentionné précédemment, la convention appliquée pour nommer les variables a été modifiée en 2007. On a enlevé l'indicateur du cycle au cours duquel les variables avaient été collectées. Cet indicateur se trouvait à la 4^e position de 2000 à 2005 (cycles 1.1 à 3.1).

Voici la liste des lettres utilisées dans les fichiers de microdonnées de l'ESCC entre 2000 et 2005 (cycles 1.1 à 3.1) et leur cycle correspondant.

Lettre	Cycle et nom du cycle
A	Cycle 1.1 : Enquête sur la santé dans les collectivités canadiennes
B	Cycle 1.2 : Enquête sur la santé dans les collectivités canadiennes, santé mentale et bien-être
C	Cycle 2.1 : Enquête sur la santé dans les collectivités canadiennes
D	Cycle 2.2 : Enquête sur la santé dans les collectivités canadiennes, nutrition
E	Cycle 3.1 : Enquête sur la santé dans les collectivités canadiennes

12.8 Lignes directrices pour l'utilisation des variables d'un sous-échantillon – Ne s'applique pas aux fichiers de 2009

12.8.2 Analyse combinant le contenu des sous-échantillons et le contenu optionnel

Le module du sous-échantillon MHW vise à fonder le calcul d'estimations au niveau national seulement. Les territoires sont exclus du sous-échantillon. Le fichier de sous-échantillon et les poids connexes ne sont pas destinés au calcul d'estimations au niveau de la région sociosanitaire.

12.9 Dictionnaires de données

Des dictionnaires de données distincts comprenant des descriptions d'univers et des fréquences sont fournis pour le fichier maître principal et le fichier de sous-échantillon.

Le traitement des modules de contenu optionnel au dictionnaire de données du fichier maître principal est identique à ce qu'il était lors des cycles antérieurs de l'ESCC. Pour chaque module, un indicateur signale si un répondant donné vit dans une région sociosanitaire où le module a été sélectionné comme contenu optionnel. Lorsque l'indicateur est 2 (non), toutes les variables du module ont des valeurs « sans objet ». Par exemple, la variable WSTFOPT indique si le module Stress au travail s'applique à un répondant donné.

12.10 Différences dans le calcul des variables de contenu commun fondé sur différents fichiers

Les variables tirées des modules de contenu commun peuvent être estimées à partir de l'un ou l'autre des deux fichiers de données lors qu'un fichier un an et un fichier deux ans sont disponibles. Selon le fichier utilisé, des différences très faibles seront observées.

Toutes les estimations officielles faites par Statistique Canada des variables des modules de contenu commun sont fondées sur le poids d'échantillonnage du fichier maître principal.

ANNEXE A

Annexe A- Enquête sur la santé dans les collectivités canadiennes (2009-20010)

Contenu commun annuel (toutes les régions sociosanitaires)					
<ul style="list-style-type: none"> • Activités physiques (PAC) • Âge du répondant (ANC) • Consommation d'alcool (ALC) • Consommation de fruits et de légumes (FVC) • Douleurs et malaise (HUP) • État de santé général (GEN) • Expériences maternelles – Allaitement (MEX) 	<ul style="list-style-type: none"> • Exposition à la fumée secondaire (ETS) • Limitation des activités (RAC) • Problèmes de santé chroniques (CCC) • Taille et poids – autodéclarés (HWT) • Usage du tabac (SMK) • Utilisation des soins de santé (HCU) • Vaccins contre la grippe (FLU) 	Administration et renseignements sociodémographiques <ul style="list-style-type: none"> • Éducation (EDU) • Personne la mieux renseignée sur la situation du ménage (PMK) • Population active (LF2) • Renseignements administratifs (ADM) • Renseignements sociodémographiques (SDC) • Revenu (INC) 			
Contenu commun deux ans (toutes les régions sociosanitaires)					
2009-2010: Blessures et santé fonctionnelle		2009 seulement: Accès aux services de soins de santéⁱ	2010 seulement: Utilisation des soins de santé et fardeau économique		
<ul style="list-style-type: none"> • Activités de tous les jours (ADL) • Blessures (INJ) • Indice de l'état de santé (HUI) • Utilisation de l'équipement protecteur (UPE) 	<ul style="list-style-type: none"> • Comportements sexuels (SXB) 	<ul style="list-style-type: none"> • Accès aux services de soins de santé (ACC) • Temps d'attente (WTM) 	<table border="0"> <tr> <td> Utilisation des soins de santé <ul style="list-style-type: none"> • Contacts avec certains professionnels de la santé (CHP) • Besoins de santé non comblés (UCN) • Les problèmes neurologiques (NEU) • Vaccin contre la grippe H1N1 (H1N) </td> <td> Fardeau économique <ul style="list-style-type: none"> • Fibromialgie (CC3) • Perte de productivité due à des raisons de santé (LOP) • Syndrome de la fatigue chronique, et multiple polysensibilité chimique (CC4) </td> </tr> </table>	Utilisation des soins de santé <ul style="list-style-type: none"> • Contacts avec certains professionnels de la santé (CHP) • Besoins de santé non comblés (UCN) • Les problèmes neurologiques (NEU) • Vaccin contre la grippe H1N1 (H1N) 	Fardeau économique <ul style="list-style-type: none"> • Fibromialgie (CC3) • Perte de productivité due à des raisons de santé (LOP) • Syndrome de la fatigue chronique, et multiple polysensibilité chimique (CC4)
Utilisation des soins de santé <ul style="list-style-type: none"> • Contacts avec certains professionnels de la santé (CHP) • Besoins de santé non comblés (UCN) • Les problèmes neurologiques (NEU) • Vaccin contre la grippe H1N1 (H1N) 	Fardeau économique <ul style="list-style-type: none"> • Fibromialgie (CC3) • Perte de productivité due à des raisons de santé (LOP) • Syndrome de la fatigue chronique, et multiple polysensibilité chimique (CC4) 				

i Demandé à un sous-échantillon de répondants. Ces modules de contenu commun un an n'ont pas été demandés aux répondants des territoires.

Contenu optionnel (certaines régions sociosanitaires)			
<ul style="list-style-type: none"> • Activités physiques – Installations au travail (PAF) • Activités sédentaires (SAC) • Auto examen des seins (BSX) • Bien-être psychologique (PWB) • Changements faits pour améliorer la santé (CIH) • Choix alimentaires (FDC) • Conduite et sécurité (DRV) • Consommation d'alcool - Dépendance (ALD) • Consommation d'alcool au cours de la dernière semaine (ALW) • Consultations au sujet de la santé mentale (CMH) • Contrôle de soi (MAS) • Couverture d'assurance (INS) • Dépistage du cancer de la prostate (PSA) • Dépistage du cancer du côlon et du rectum (CCS) 	<ul style="list-style-type: none"> • Dépression (DEP) • Détresse (DIS) • Estime de soi (SFE) • État de santé (SF-36) (SFR) • Examen des seins (BRX) • Examens de la vue (EYX) • Expériences maternelles - Consommation d'alcool au cours de la grossesse (MXA) • Expériences maternelles – Usage du tabac au cours de la grossesse (MXS) • Jeu excessif (CPG) • Mammographie(MAM) • Mesures de sécurité à la maison (HMS) • Organismes à but non lucratif - Participation (ORG) • Pensées suicidaires et tentatives de suicide (SUI) • Protection contre le soleil (SSB) • Santé bucco-dentaire 1 (OH2) • Satisfaction à l'égard de la vie (SWL) • Satisfaction à l'égard du système de soins de santé (HCS) 	<ul style="list-style-type: none"> • Satisfaction des patients à l'égard des services de soins de santé (PAS) • Satisfaction des patients à l'égard des soins de santé communautaires (PSC) • Sécurité alimentaire (FSC) • Services de soins de santé à domicile (HMC) • Soins pour le diabète (DIA) • Soutien social - Disponibilité (SSA) • Soutien social - Utilisation (SSU) • Stress - Faire face au stress (STC) • Stress – Sources (STS) • Tension artérielle - Vérification (BPC) • Test pap (PAP) • Usage de drogues illicites (IDG) 	<ul style="list-style-type: none"> • Usage de suppléments alimentaires - Vitamines et minéraux (DSU) • Usage du tabac - Autres produits du tabac (TAL) • Usage du tabac - Consultation d'un médecin (SPC) • Usage du tabac - Étapes du changement (SCH) • Usage du tabac - Méthodes pour cesser de fumer (SCA) • Visites chez le dentiste (DEN)
Réponse rapide			
2009		2010	
<ul style="list-style-type: none"> • Apnée du sommeil (SLA) (jan. – fév. 2009) • Ostéoporose (OST) (mar. – avr. 2009) • Infertilité (IFT) (sep. – déc. 2009) 		<ul style="list-style-type: none"> • Stigma et discrimination de la santé mentale (STG) (mai – juin 2010) • Expérience de santé mentale (MHE) (mai – juin 2010) 	

ANNEXE B

Annexe B – Sélection du contenu optionnel par les provinces et territoires

Modules optionnels	T.-N.-L.	Î.-P.-É.	N.-É.	N.-B.	Qc	Ont.	Man.	Sask.	Alb.	C.-B.	Yn	T.N.-O.	Nt
Consommation d'alcool – Dépendance (ALD)													•
Consommation d'alcool au cours de la dernière semaine (ALW)	•					•		•					
Tension artérielle – Vérification (BPC)		•		•							•	•	
Examen des seins (BRX)		•						•			•	•	•
Auto-examen des seins (BSX)											•		•
Changements faits pour améliorer la santé (CIH)		•	•							•	•		
Dépistage du cancer du côlon et du rectum (CCS)	•	•	•	•		•		•			•	•	•
Consultations au sujet de la santé mentale (CMH)		•		•	•	•	•	•	•		•	•	
Visites chez le dentiste (DEN)		•	•			•					•		
Dépression (DEP)		•			•			•	•	•		•	•
Soins pour le diabète (DIA)	•	•											
Usage de suppléments alimentaires - Vitamines et minéraux (DSU)												•	
Détresse (DIS)					•			•	•			•	
Conduite et sécurité (DRV)	•					•			•		•		
Examens de la vue (EYX)			•			•					•		•
Choix alimentaires (FDC)			•		•						•	•	
Sécurité alimentaire (FSC)	•		•		•	•	•	•	•	•	•	•	•
Satisfaction à l'égard du système de soins de santé (HCS)	•					•		•				•	
Services de soins de santé à domicile (HMC)						•							
Mesures de sécurité à la maison (HMS)													•

Modules optionnels	T.-N.-L.	Î.-P.-É.	N.-É.	N.-B.	Qc	Ont.	Man.	Sask.	Alb.	C.-B.	Yn	T.N.-O.	Nt
Usage de drogues illicites (IDG)				•		•		•					•
Couverture d'assurance (INS)													•
Mammographie (MAM)	•		•	•					•			•	
Contrôle de soi (MAS)							•					•	
Expériences maternelles – Consommation d'alcool au cours de la grossesse (MXA)									•		•		•
Expériences maternelles – Usage du tabac au cours de la grossesse (MXS)						•			•		•		•
Santé bucco-dentaire 2 (OH2)		•				•							
Test pap (PAP)		•	•								•		•
Satisfaction des patients à l'égard des services de soins de santé (PAS)						•							
Satisfaction des patients à l'égard des soins de santé communautaires (PSC)			•						•				•
Activités physiques – Installations au travail (PAF)			•										
Jeu excessif (CPG)				•					•				•
Dépistage du cancer de la prostate (PSA)	•	•	•								•	•	
Bien-être psychologique (PWB)				•									
Satisfaction à l'égard de la vie (SWL)										•			
Activités sédentaires (SAC)	•						•			•			
Estime de soi (SFE)					•					•		•	
État de santé (SF-36) (SFR)							•						
Usage du tabac – Consultation d'un médecin (SPC)							•				•		•
Usage du tabac – Les étapes du changement (SCH)							•						

Modules optionnels	T.-N.-L.	Î.-P.-É.	N.-É.	N.-B.	Qc	Ont.	Man.	Sask.	Alb.	C.-B.	Yn	T.N.-O.	Nt
Usage du tabac - Méthodes pour cesser de fumer (SCA)											•		•
Soutien social – Disponibilité (SSA)				•	•			•		•		•	
Soutien social – Utilisation (SSU)				•						•			
Stress - Faire face au stress (STC)											•		
Stress – Sources (STS)			•								•		
Pensées suicidaires et tentatives de suicide (SUI)	•								•	•			•
Protection contre le soleil (SSB)		•	•		•								
Usage du tabac - Autres produits du tabac (TAL)						•							
Organismes à but non lucratif – Participation (ORG)										•		•	

ANNEXE C

Annexe C - Géographie disponible sur le fichier maître et de partage et codes correspondants : Canada, provinces/territoires, régions sociosanitaires et groupes homologues

0	Canada	
10	Terre-Neuve-et-Labrador	
1011-C		Eastern Regional Integrated Health Authority
1012-I		Central Regional Integrated Health Authority
1013-I		Western Regional Integrated Health Authority
1014-H		Labrador-Grenfell Regional Integrated Health Authority
11	Île-du-Prince-Édouard	
1101-D		Kings County
1102-A		Queens County
1103-C		Prince County
12	Nouvelle-Écosse	
1201-C		Zone 1
1202-C		Zone 2
1203-C		Zone 3
1204-C		Zone 4
1205-I		Zone 5
1206-A		Zone 6
13	Nouveau-Brunswick	
1301-C		Région 1
1302-C		Région 2
1303-C		Région 3
1304-C		Région 4
1305-I		Région 5
1306-I		Région 6
1307-I		Région 7
24	Québec	
2401-C		Région du Bas-Saint-Laurent
2402-C		Région du Saguenay - Lac-Saint-Jean
2403-A		Région de la Capitale-Nationale
2404-C		Région de la Mauricie et du Centre-du-Québec
2405-C		Région de l'Estrie
2406-G		Région de Montréal
2407-A		Région de l'Outaouais
2408-C		Région de l'Abitibi-Témiscamingue
2409-H		Région de la Côte-Nord
2410-H		Région du Nord-du-Québec
2411-I		Région de la Gaspésie - Îles-de-la-Madeleine
2412-E		Région de la Chaudière-Appalaches
2413-A		Région de Laval
2414-E		Région de Lanaudière
2415-E		Région des Laurentides
2416-A		Région de la Montérégie

35	Ontario par Réseau local d'intégration des services de santé
3501	Réseau d'intégration des services de santé de Érié St. Clair
3502	Réseau d'intégration des services de santé du Sud-Ouest
3503	Réseau d'intégration des services de santé de Waterloo Wellington
3504	Réseau d'intégration des services de santé de Hamilton Niagara Haldimand Brant
3505	Réseau d'intégration des services de santé du Centre-Ouest
3506	Réseau d'intégration des services de santé de Mississauga Halton
3507	Réseau d'intégration des services de santé de Toronto-Centre
3508	Réseau d'intégration des services de santé du Centre
3509	Réseau d'intégration des services de santé du Centre-Est
3510	Réseau d'intégration des services de santé du Sud-Est
3511	Réseau d'intégration des services de santé de Champlain
3512	Réseau d'intégration des services de santé de Simcoe-Nord Muskoka
3513	Réseau d'intégration des services de santé du Nord-Est
3514	Réseau d'intégration des services de santé du Nord-Ouest
35	Ontario par circonscription sanitaire
3526-C	Circonscription sanitaire du district d'Algoma
3527-A	Circonscription sanitaire du comté de Brant
3530-B	Circonscription sanitaire régionale de Durham
3531-E	Circonscription sanitaire d'Elgin-St. Thomas
3533-E	Circonscription sanitaire de Grey Bruce
3534-E	Circonscription sanitaire de Haldimand-Norfolk
3535-E	Circonscription sanitaire du district de Haliburton, Kawartha et Pine Ridge
3536-B	Circonscription sanitaire régionale de Halton
3537-A	Circonscription sanitaire de la cité de Hamilton
3538-A	Circonscription sanitaire des comtés de Hastings et Prince Edward
3539-E	Circonscription sanitaire du comté de Huron
3540-A	Circonscription sanitaire de Chatham-Kent
3541-A	Circonscription sanitaire de Kingston, Frontenac et Lennox et Addington
3542-A	Circonscription sanitaire de Lambton
3543-E	Circonscription sanitaire de Leeds, Grenville et Lanark
3544-A	Circonscription sanitaire de Middlesex-London
3546-A	Circonscription sanitaire régionale de Niagara
3547-C	Circonscription sanitaire du district de North Bay Parry Sound
3549-H	Circonscription sanitaire du Nord-Ouest
3551-B	Circonscription sanitaire de la cité d'Ottawa
3552-E	Circonscription sanitaire du comté d'Oxford
3553-B	Circonscription sanitaire régionale de Peel
3554-E	Circonscription sanitaire du district de Perth
3555-A	Circonscription sanitaire du comté et de la cité de Peterborough
3556-H	Circonscription sanitaire de Porcupine
3557-E	Circonscription sanitaire du comté et du district de Renfrew
3558-E	Circonscription sanitaire de l'Est de l'Ontario
3560-E	Circonscription sanitaire du district de Simcoe Muskoka
3561-C	Circonscription sanitaire de Sudbury et son district
3562-C	Circonscription sanitaire du district de Thunder Bay
3563-C	Circonscription sanitaire de Timiskaming
3565-B	Circonscription sanitaire de Waterloo
3566-B	Circonscription sanitaire de Wellington-Dufferin-Guelph
3568-B	Circonscription sanitaire de Windsor-Comté d'Essex

3570-B		Circonscription sanitaire de York
3595-G		Circonscription sanitaire de la cité de Toronto
46	Manitoba	
4610-A		Winnipeg Regional Health Authority
4615-A		Brandon Regional Health Authority
4620-E		North Eastman Regional Health Authority
4625-E		South Eastman Regional Health Authority
4630-E		Interlake Regional Health Authority
4640-D		Central Regional Health Authority
4645-D		Assiniboine Regional Health Authority
4660-D		Parkland Regional Health Authority
4670-H		Nor-Man Regional Health Authority
4685-F		Burntwood/Churchill
47	Saskatchewan	
4701-D		Sun Country Regional Health Authority
4702-D		Five Hills Regional Health Authority
4703-D		Cypress Regional Health Authority
4704-A		Regina Qu'Appelle Regional Health Authority
4705-D		Sunrise Regional Health Authority
4706-A		Saskatoon Regional Health Authority
4707-D		Heartland Regional Health Authority
4708-D		Kelsey Trail Regional Health Authority
4709-C		Prince Albert Parkland Regional Health Authority
4710-H		Prairie North Regional Health Authority
4714-F		Mamawetan/Keewatin/Athabasca
48	Alberta	
4821-E		Chinook Regional Health Authority
4822-E		Palliser Health Region
4823-B		Calgary Health Region
4824-E		David Thompson Regional Health Authority
4825-E		East Central Health
4826-E		Capital Health
4827-E		Aspen Regional Health Authority
4828-E		Peace Country Health
4829-H		Northern Lights Health Region
59	Colombie-Britannique	
5911-E		East Kootenay Health Service Delivery Area
5912-C		Kootenay-Boundary Health Service Delivery Area
5913-A		Okanagan Health Service Delivery Area
5914-C		Thompson/Cariboo Health Service Delivery Area
5921-A		Fraser East Health Service Delivery Area
5922-B		Fraser North Health Service Delivery Area
5923-B		Fraser South Health Service Delivery Area
5931-B		Richmond Health Service Delivery Area
5932-G		Vancouver Health Service Delivery Area
5933-B		North Shore/Coast Garibaldi Health Service Delivery Area
5941-A		South Vancouver Island Health Service Delivery Area
5942-A		Central Vancouver Island Health Service Delivery Area
5943-C		North Vancouver Island Health Service Delivery Area

5951-H		Northwest Health Service Delivery Area
5952-H		Northern Interior Health Service Delivery Area
5953-H		Northeast Health Service Delivery Area
60	Yukon	
6001-H		Yukon
61	Territoires du Nord-Ouest	
6101-H		Territoires du Nord-Ouest
62	Nunavut – 10 communautés les plus grandes	
6201-F		Nunavut – 10 communautés les plus grandes
A	Groupe de régions homologues A	
B	Groupe de régions homologues B	
C	Groupe de régions homologues C	
D	Groupe de régions homologues D	
E	Groupe de régions homologues E	
F	Groupe de régions homologues F	
G	Groupe de régions homologues G	
H	Groupe de régions homologues H	
I	Groupe de régions homologues I	

ANNEXE D

Annexe D - Répartition de l'échantillon par région sociosanitaire et par base de sondage de l'ESCC 2009

Prov./Terr.	Région sociosanitaire	Base aréolaire		Bases téléphoniques		Combiné	
		# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute
Canada	Total	33 136	47 888	32 626	55 920	65 762	103 808
T.-N.-L.	Total	1 003	1 404	1 002	1 596	2 005	3 000
	1 011	405	575	405	648	810	1 223
	1 012	235	333	235	360	470	693
	1 013	213	282	212	342	425	624
	1 014	150	214	150	246	300	460
I.-P.-É.	Total	501	785	500	948	1 001	1 733
	1 101	89	159	89	168	178	327
	1102	230	366	230	420	460	786
	1103	182	260	181	360	363	620
N.-É.	Total	1 261	1 903	1 259	1 920	2 520	3 823
	1201	198	306	197	312	395	618
	1202	160	236	160	252	320	488
	1203	180	239	180	270	360	509
	1204	175	321	175	270	350	591
	1205	210	284	210	324	420	608
	1206	338	517	337	492	675	1 009
N.-B.	Total	1 289	1 940	1 286	1 938	2 575	3 878
	1301	250	387	250	384	500	771
	1302	243	403	242	378	485	781
	1303	235	373	235	366	470	739
	1304	135	202	135	198	270	400
	1305	125	170	125	180	250	350
	1306	173	239	172	240	345	479
	1307	128	166	127	192	255	358
Qc	Total	5 874	8 104	6 270	10 998	12 144	19 102
	2401	300	383	300	474	600	857
	2402	314	423	314	546	628	969
	2403	463	647	463	726	926	1 373
	2404	402	521	401	612	803	1 133
	2405	309	455	309	474	618	929
	2406	777	1 114	776	1 458	1 553	2 572
	2407	323	508	322	522	645	1 030
	2408	300	383	300	456	600	839
	2409	300	405	300	558	600	963
	2410	0	0	400	1 248	400	1 248
	2411	300	411	300	516	600	927
	2412	362	484	361	636	723	1 120
	2413	335	480	335	564	670	1 044
	2414	359	480	359	588	718	1 068
	2415	380	536	380	624	760	1 160
	2416	650	875	650	996	1 300	1 871

Prov./Terr.	Région socio-sanitaire	Base aréolaire		Bases téléphoniques		Combiné	
		# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute
Ont.	Total	11 111	15 896	11 096	19 158	22 207	35 054
	3526	213	318	212	336	425	654
	3527	203	282	202	312	405	594
	3530	408	574	407	648	815	1 222
	3531	170	239	170	276	340	515
	3533	240	359	252	450	492	809
	3534	193	290	193	348	386	638
	3535	238	316	237	414	475	730
	3536	353	499	352	552	705	1 051
	3537	413	622	412	696	825	1 318
	3538	235	317	235	438	470	755
	3539	148	228	147	276	295	504
	3540	200	250	200	330	400	580
	3541	253	380	252	450	505	830
	3542	218	286	217	366	435	652
	3543	238	335	237	378	475	713
	3544	375	564	375	630	750	1 194
	3546	383	505	382	618	765	1 123
	3547	200	310	200	384	400	694
	3549	200	335	200	438	400	773
	3551	513	750	512	810	1 025	1 560
	3552	188	245	187	282	375	527
	3553	671	949	670	1 206	1 341	2 155
	3554	163	223	162	246	325	469
	3555	213	316	212	384	425	700
	3556	188	286	187	264	375	550
	3557	188	278	187	324	375	602
	3558	260	355	260	396	520	751
	3560	560	805	560	1 074	1 120	1 879
	3561	270	393	270	480	540	873
	3562	332	483	331	588	663	1 071
	3563	125	183	125	216	250	399
	3565	383	560	382	600	765	1 160
	3566	293	382	292	456	585	838
	3568	358	497	357	600	715	1 097
	3570	473	627	472	870	945	1 497
	3595	1 052	1 554	1 048	2 022	2 100	3 576
Man.	Total	1 877	2 584	1 873	2 988	3 750	5 572
	4610	528	714	527	786	1 055	1 500
	4615	140	200	140	228	280	428
	4620	125	168	125	228	250	396
	4625	150	205	150	240	300	445
	4630	173	272	172	282	345	554
	4640	200	253	200	276	400	529
	4645	178	239	177	270	355	509

Prov./Terr.	Région socio-sanitaire	Base aréolaire		Bases téléphoniques		Combiné	
		# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute
	4660	133	190	132	210	265	400
	4670	125	182	125	228	250	410
	4685	125	160	125	240	250	400
Sask.	Total	1 806	2 555	2 054	4 098	3 860	6 653
	4701	150	187	150	228	300	415
	4702	150	201	150	264	300	465
	4703	133	185	132	222	265	407
	4704	310	434	310	504	620	938
	4705	155	216	155	228	310	444
	4706	330	455	330	528	660	983
	4707	135	213	135	216	270	429
	4708	130	189	130	204	260	393
	4709	163	286	162	306	325	592
	4710	150	189	150	258	300	447
	4714	0	0	250	1 140	250	1 140
Alb.	Total	3 052	4 490	3 048	5 076	6 100	9 566
	4821	255	373	255	408	510	781
	4822	208	259	207	318	415	577
	4823	698	996	697	1 122	1 395	2 118
	4824	350	513	350	588	700	1 101
	4825	223	290	222	336	445	626
	4826	655	988	655	1 146	1 310	2 134
	4827	270	404	270	462	540	866
	4828	233	376	232	396	465	772
	4829	160	292	160	300	320	592
C.-B.	Total	4 027	6 092	4 023	6 612	8 050	12 704
	5911	153	232	152	264	305	496
	5912	155	236	155	234	310	470
	5913	295	394	295	486	590	880
	5914	250	332	250	402	500	734
	5921	260	353	260	402	520	755
	5922	380	543	380	588	760	1 131
	5923	400	605	400	690	800	1 295
	5931	213	287	213	336	426	623
	5932	400	639	400	714	800	1 353
	5933	273	496	273	474	546	970
	5941	338	495	338	528	676	1 023
	5942	263	373	263	384	526	757
	5943	133	181	132	210	265	391
	5951	163	311	163	318	326	629
	5952	213	307	213	360	426	667
	5903	138	306	138	222	276	528
Yn	6001	475	738	125	306	600	1 044
T.N.-O.	6101	510	816	90	282	600	1 098
Nt	6201	350	580	0	0	350	580

**Répartition de l'échantillon par réseau local d'intégration d'intégration des services de santé
et par base de sondage de l'ESCC 2009 en Ontario**

		Base aréolaire		Bases téléphoniques		Combiné	
Prov./Terr.	RLISS	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute
Ont.	Total	11 111	15 896	11 096	19 158	22 207	35 054
	3501	776	1 033	774	1 296	1 550	2 329
	3502	1 282	1 853	1 279	2 130	2 561	3 983
	3503	622	867	620	978	1 242	1 845
	3504	1 300	1 850	1 297	2 136	2 597	3 986
	3505	536	772	533	954	1 069	1 726
	3506	558	772	557	942	1 115	1 714
	3507	541	855	540	1 044	1 081	1 899
	3508	706	940	705	1 320	1 411	2 260
	3509	1 056	1 469	1 052	1 830	2 108	3 299
	3510	657	941	656	1 164	1 313	2 105
	3511	1 030	1 474	1 027	1 632	2 057	3 106
	3512	519	759	531	1 026	1 050	1 785
	3513	996	1 491	994	1 680	1 990	3 171
	3514	532	819	531	1 026	1 063	1 845

**Répartition de l'échantillon par province et par base de sondage du sous-échantillon de
l'Enquête sur l'accès aux services de soins de santé (EASS) 2009¹**

		Base aréolaire		Bases téléphoniques		Combiné	
Prov./Terr.	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute	# attendu de répondants	taille d'échantillon brute	
Canada	23 593	35 983	23 571	42 309	47 164	78 292	
T.-N.-L.	940	1 404	939	1 596	1 879	3 000	
Î.-P.-É.	470	785	469	948	939	1 733	
N.-É.	1 188	1 903	1 186	1 920	2 374	3 823	
N.-B.	1 198	1 940	1 195	1 938	2 393	3 878	
Qc	2 300	3 379	2 300	4 097	4 600	7 476	
Ont.	10 497	15 896	10 482	19 158	20 979	35 054	
Man.	1 600	2 319	1 600	2 654	3 200	4 974	
Sask.	1 600	2 336	1 600	3 262	3 200	5 598	
Alb.	1 800	2 806	1 800	3 190	3 600	5 996	
C.-B.	2 000	3 215	2 000	3 545	4 000	6 760	

¹Les répondants de l'ESCC âgés de 12 à 14 ans et les répondants par personne interposée sont considérés hors du champ de l'enquête pour l'EASS. Ceci explique les différences observées pour certaines provinces entre les comptes du présent tableau et ceux présentés à la section 5.8. Pour les provinces où toutes les unités de l'ESCC ont été sélectionnées pour l'EASS, le nombre attendu de répondants a été ajusté dans le présent tableau pour tenir compte du nombre attendu d'unités qui devraient être hors du champ de l'enquête pour l'EASS. Pour les autres provinces, il a été possible de tenir compte de la présence de telles unités au moment de l'échantillonnage, en sélectionnant une plus grande proportion des répondants de l'ESCC, de sorte à obtenir le nombre attendu de répondants désiré pour l'EASS.

ANNEXE E

Annexe E - Taux de réponse par région sociosanitaire et par base de sondage de l'ESCC 2009

Prov. Terr.	Région socio- sanitaire	Base aéroilaire							Bases téléphoniques							
		# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné
Canada	Total	40 136	33 307	83,0	33 307	30 475	91,5	75,9	44 125	35 219	79,8	35 219	31 204	88,6	70,7	73,2
T.-N.-L.	Total	1 141	991	86,9	991	906	91,4	79,4	1 336	1 137	85,1	1 137	992	87,2	74,3	76,6
	1011	469	386	82,3	386	342	88,6	72,9	548	467	85,2	467	402	86,1	73,4	73,2
	1012	251	229	91,2	229	210	91,7	83,7	306	259	84,6	259	227	87,6	74,2	78,5
	1013	235	217	92,3	217	204	94,0	86,8	278	232	83,5	232	209	90,1	75,2	80,5
	1014	186	159	85,5	159	150	94,3	80,6	204	179	87,7	179	154	86,0	75,5	77,9
Î.-P.-É.	Total	620	524	84,5	524	470	89,7	75,8	645	545	84,5	545	470	86,2	72,9	74,3
	1101	103	90	87,4	90	84	93,3	81,6	55	47	85,5	47	40	85,1	72,7	78,5
	1102	308	261	84,7	261	232	88,9	75,3	338	286	84,6	286	245	85,7	72,5	73,8
	1103	209	173	82,8	173	154	89,0	73,7	252	212	84,1	212	185	87,3	73,4	73,5
N.-É.	Total	1 531	1 308	85,4	1 308	1 181	90,3	77,1	1 560	1 338	85,8	1 338	1 189	88,9	76,2	76,7
	1201	223	214	96,0	214	199	93,0	89,2	240	206	85,8	206	188	91,3	78,3	83,6
	1202	189	161	85,2	161	148	91,9	78,3	205	176	85,9	176	158	89,8	77,1	77,7
	1203	186	156	83,9	156	148	94,9	79,6	197	165	83,8	165	149	90,3	75,6	77,5
	1204	247	220	89,1	220	203	92,3	82,2	220	192	87,3	192	169	88,0	76,8	79,7
	1205	240	197	82,1	197	179	90,9	74,6	258	213	82,6	213	185	86,9	71,7	73,1
	1206	446	360	80,7	360	304	84,4	68,2	440	386	87,7	386	340	88,1	77,3	72,7
N.-B.	Total	1 493	1 283	85,9	1 283	1 154	89,9	77,3	1 605	1 387	86,4	1 387	1 252	90,3	78,0	77,7
	1301	321	264	82,2	264	239	90,5	74,5	327	284	86,9	284	255	89,8	78,0	76,2
	1302	290	242	83,4	242	226	93,4	77,9	312	272	87,2	272	237	87,1	76,0	76,9
	1303	254	222	87,4	222	199	89,6	78,3	310	268	86,5	268	251	93,7	81,0	79,8
	1304	167	142	85,0	142	128	90,1	76,6	157	140	89,2	140	127	90,7	80,9	78,7
	1305	129	116	89,9	116	95	81,9	73,6	149	133	89,3	133	122	91,7	81,9	78,1
	1306	203	189	93,1	189	174	92,1	85,7	197	168	85,3	168	149	88,7	75,6	80,8

		Base aréolaire							Bases téléphoniques							
Prov. Terr.	Région socio-sanitaire	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné
	1307	129	108	83,7	108	93	86,1	72,1	153	122	79,7	122	111	91,0	72,5	72,3
Qc	Total	7 087	5 789	81,7	5 789	5 412	93,5	76,4	8 767	7 003	79,9	7 003	6 170	88,1	70,4	73,1
	2401	314	290	92,4	290	277	95,5	88,2	375	316	84,3	316	279	88,3	74,4	80,7
	2402	365	305	83,6	305	286	93,8	78,4	457	394	86,2	394	359	91,1	78,6	78,5
	2403	608	473	77,8	473	451	95,3	74,2	627	503	80,2	503	455	90,5	72,6	73,4
	2404	457	389	85,1	389	364	93,6	79,6	526	432	82,1	432	393	91,0	74,7	77,0
	2405	354	265	74,9	265	254	95,8	71,8	391	331	84,7	331	299	90,3	76,5	74,2
	2406	1 001	735	73,4	735	684	93,1	68,3	1 258	922	73,3	922	770	83,5	61,2	64,4
	2407	417	338	81,1	338	310	91,7	74,3	438	358	81,7	358	323	90,2	73,7	74,0
	2408	319	281	88,1	281	262	93,2	82,1	398	329	82,7	329	293	89,1	73,6	77,4
	2409	346	300	86,7	300	290	96,7	83,8	391	300	76,7	300	257	85,7	65,7	74,2
	2410	562	450	80,1	450	400	88,9	71,2	71,2
	2411	341	315	92,4	315	298	94,6	87,4	393	309	78,6	309	272	88,0	69,2	77,7
	2412	435	389	89,4	389	367	94,3	84,4	541	422	78,0	422	375	88,9	69,3	76,0
	2413	428	334	78,0	334	308	92,2	72,0	500	394	78,8	394	342	86,8	68,4	70,0
	2414	433	359	82,9	359	329	91,6	76,0	511	410	80,2	410	356	86,8	69,7	72,6
	2415	475	380	80,0	380	339	89,2	71,4	497	397	79,9	397	353	88,9	71,0	71,2
	2416	794	636	80,1	636	593	93,2	74,7	902	736	81,6	736	644	87,5	71,4	72,9
Ont.	Total	13 662	11 229	82,2	11 229	10 211	90,9	74,7	15 703	12 256	78,0	12 256	10 758	87,8	68,5	71,4
	3526	287	253	88,2	253	233	92,1	81,2	264	211	79,9	211	193	91,5	73,1	77,3
	3527	266	219	82,3	219	183	83,6	68,8	252	199	79,0	199	179	89,9	71,0	69,9
	3530	523	427	81,6	427	388	90,9	74,2	566	447	79,0	447	376	84,1	66,4	70,2
	3531	203	166	81,8	166	145	87,3	71,4	234	189	80,8	189	165	87,3	70,5	70,9
	3533	306	281	91,8	281	264	94,0	86,3	334	261	78,1	261	234	89,7	70,1	77,8
	3534	247	195	78,9	195	179	91,8	72,5	281	216	76,9	216	188	87,0	66,9	69,5
	3535	216	175	81,0	175	152	86,9	70,4	279	237	84,9	237	208	87,8	74,6	72,7

	3536	476	390	81,9	390	359	92,1	75,4	488	382	78,3	382	334	87,4	68,4	71,9
--	------	-----	-----	------	-----	-----	------	------	-----	-----	------	-----	-----	------	------	------

		Base aréolaire							Bases téléphoniques								
Prov. Terr.	Région socio-sanitaire	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné	
	3537	547	420	76,8	420	373	88,8	68,2	590	459	77,8	459	403	87,8	68,3	68,2	
	3538	276	234	84,8	234	212	90,6	76,8	347	278	80,1	278	242	87,1	69,7	72,9	
	3539	193	175	90,7	175	167	95,4	86,5	226	188	83,2	188	167	88,8	73,9	79,7	
	3540	209	198	94,7	198	191	96,5	91,4	247	199	80,6	199	177	88,9	71,7	80,7	
	3541	334	268	80,2	268	235	87,7	70,4	339	270	79,6	270	243	90,0	71,7	71,0	
	3542	245	203	82,9	203	187	92,1	76,3	296	249	84,1	249	222	89,2	75,0	75,6	
	3543	264	220	83,3	220	195	88,6	73,9	319	255	79,9	255	223	87,5	69,9	71,7	
	3544	481	380	79,0	380	359	94,5	74,6	527	408	77,4	408	367	90,0	69,6	72,0	
	3546	448	373	83,3	373	338	90,6	75,4	520	391	75,2	391	352	90,0	67,7	71,3	
	3547	251	215	85,7	215	188	87,4	74,9	270	216	80,0	216	183	84,7	67,8	71,2	
	3549	256	186	72,7	186	168	90,3	65,6	298	241	80,9	241	215	89,2	72,1	69,1	
	3551	687	489	71,2	489	442	90,4	64,3	705	564	80,0	564	498	88,3	70,6	67,5	
	3552	228	200	87,7	200	194	97,0	85,1	257	208	80,9	208	184	88,5	71,6	77,9	
	3553	891	758	85,1	758	677	89,3	76,0	1 060	825	77,8	825	705	85,5	66,5	70,8	
	3554	211	190	90,0	190	183	96,3	86,7	212	174	82,1	174	155	89,1	73,1	79,9	
	3555	243	206	84,8	206	186	90,3	76,5	284	225	79,2	225	209	92,9	73,6	75,0	
	3556	258	210	81,4	210	184	87,6	71,3	225	168	74,7	168	148	88,1	65,8	68,7	
	3557	211	203	96,2	203	184	90,6	87,2	257	197	76,7	197	174	88,3	67,7	76,5	
	3558	306	248	81,0	248	232	93,5	75,8	336	259	77,1	259	232	89,6	69,0	72,3	
	3560	650	509	78,3	509	467	91,7	71,8	786	616	78,4	616	553	89,8	70,4	71,0	
	3561	302	267	88,4	267	230	86,1	76,2	381	297	78,0	297	279	93,9	73,2	74,5	
	3562	394	312	79,2	312	289	92,6	73,4	430	342	79,5	342	308	90,1	71,6	72,5	
	3563	13	11	84,6	11	10	90,9	76,9	181	141	77,9	141	119	84,4	65,7	66,5	
	3565	493	408	82,8	408	372	91,2	75,5	521	409	78,5	409	364	89,0	69,9	72,6	
	3566	344	308	89,5	308	290	94,2	84,3	370	303	81,9	303	267	88,1	72,2	78,0	
	3568	431	358	83,1	358	333	93,0	77,3	533	406	76,2	406	348	85,7	65,3	70,6	

Prov. Terr.	Région socio- sanitaire	Base aréolaire							Bases téléphoniques							
		# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné
	3570	572	470	82,2	470	430	91,5	75,2	758	589	77,7	589	506	85,9	66,8	70,4
	3595	1 400	1 104	78,9	1 104	992	89,9	70,9	1 730	1 237	71,5	1 237	1 038	83,9	60,0	64,9
Man.	Total	2 195	1 849	84,2	1 849	1 676	90,6	76,4	2 212	1 859	84,0	1 859	1 693	91,1	76,5	76,4
	4610	657	520	79,1	520	463	89,0	70,5	686	575	83,8	575	522	90,8	76,1	73,3
	4615	189	148	78,3	148	131	88,5	69,3	185	153	82,7	153	139	90,8	75,1	72,2
	4620	138	125	90,6	125	117	93,6	84,8	122	105	86,1	105	97	92,4	79,5	82,3
	4625	177	147	83,1	147	131	89,1	74,0	182	161	88,5	161	143	88,8	78,6	76,3
	4630	189	164	86,8	164	149	90,9	78,8	189	165	87,3	165	154	93,3	81,5	80,2
	4640	237	212	89,5	212	196	92,5	82,7	216	177	81,9	177	156	88,1	72,2	77,7
	4645	212	191	90,1	191	168	88,0	79,2	220	177	80,5	177	163	92,1	74,1	76,6
	4660	129	110	85,3	110	105	95,5	81,4	165	131	79,4	131	120	91,6	72,7	76,5
	4670	145	127	87,6	127	121	95,3	83,4	148	127	85,8	127	117	92,1	79,1	81,2
	4685	122	105	86,1	105	95	90,5	77,9	99	88	88,9	88	82	93,2	82,8	80,1
Sask.	Total	2 074	1 845	89,0	1 845	1 749	94,8	84,3	2 742	2 247	81,9	2 247	2 051	91,3	74,8	78,9
	4701	156	152	97,4	152	148	97,4	94,9	182	151	83,0	151	135	89,4	74,2	83,7
	4702	171	150	87,7	150	142	94,7	83,0	222	179	80,6	179	165	92,2	74,3	78,1
	4703	142	122	85,9	122	120	98,4	84,5	182	154	84,6	154	138	89,6	75,8	79,6
	4704	385	354	91,9	354	323	91,2	83,9	427	340	79,6	340	312	91,8	73,1	78,2
	4705	155	140	90,3	140	133	95,0	85,8	179	145	81,0	145	132	91,0	73,7	79,3
	4706	390	328	84,1	328	313	95,4	80,3	466	387	83,0	387	348	89,9	74,7	77,2
	4707	138	121	87,7	121	118	97,5	85,5	167	135	80,8	135	128	94,8	76,6	80,7
	4708	145	131	90,3	131	125	95,4	86,2	161	141	87,6	141	130	92,2	80,7	83,3
	4709	243	208	85,6	208	198	95,2	81,5	219	184	84,0	184	173	94,0	79,0	80,3
	4710	149	139	93,3	139	129	92,8	86,6	171	141	82,5	141	130	92,2	76,0	80,9
	4714	366	290	79,2	290	260	89,7	71,0	71,0
Alb.	Total	3 743	3 037	81,1	3 037	2 709	89,2	72,4	4 068	3 238	79,6	3 238	2 900	89,6	71,3	71,8

	4821	304	246	80,9	246	232	94,3	76,3	323	273	84,5	273	248	90,8	76,8	76,6
--	-------------	-----	-----	------	-----	-----	------	------	-----	-----	------	-----	-----	------	------	------

		Base aréolaire							Bases téléphoniques								
Prov. Terr.	Région socio-sanitaire	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné	
	4822	223	190	85,2	190	180	94,7	80,7	225	183	81,3	183	163	89,1	72,4	76,6	
	4823	852	686	80,5	686	634	92,4	74,4	942	747	79,3	747	663	88,8	70,4	72,3	
	4824	427	351	82,2	351	313	89,2	73,3	448	375	83,7	375	341	90,9	76,1	74,7	
	4825	247	191	77,3	191	175	91,6	70,9	269	210	78,1	210	186	88,6	69,1	70,0	
	4826	837	645	77,1	645	531	82,3	63,4	952	743	78,0	743	662	89,1	69,5	66,7	
	4827	326	272	83,4	272	253	93,0	77,6	357	279	78,2	279	255	91,4	71,4	74,4	
	4828	323	292	90,4	292	252	86,3	78,0	319	254	79,6	254	230	90,6	72,1	75,1	
	4829	204	164	80,4	164	139	84,8	68,1	233	174	74,7	174	152	87,4	65,2	66,6	
C.-B.	Total	5 023	4 072	81,1	4 072	3 725	91,5	74,2	5 269	4 029	76,5	4 029	3 562	88,4	67,6	70,8	
	5911	194	164	84,5	164	151	92,1	77,8	206	157	76,2	157	148	94,3	71,8	74,8	
	5912	166	152	91,6	152	142	93,4	85,5	186	145	78,0	145	132	91,0	71,0	77,8	
	5913	345	303	87,8	303	287	94,7	83,2	407	314	77,1	314	285	90,8	70,0	76,1	
	5914	278	243	87,4	243	226	93,0	81,3	293	235	80,2	235	212	90,2	72,4	76,7	
	5921	304	255	83,9	255	231	90,6	76,0	304	243	79,9	243	207	85,2	68,1	72,0	
	5922	473	383	81,0	383	362	94,5	76,5	505	371	73,5	371	330	88,9	65,3	70,8	
	5923	535	448	83,7	448	408	91,1	76,3	575	440	76,5	440	386	87,7	67,1	71,5	
	5931	232	198	85,3	198	180	90,9	77,6	282	205	72,7	205	170	82,9	60,3	68,1	
	5932	527	382	72,5	382	365	95,5	69,3	570	385	67,5	385	321	83,4	56,3	62,5	
	5933	332	270	81,3	270	209	77,4	63,0	380	275	72,4	275	246	89,5	64,7	63,9	
	5941	429	336	78,3	336	312	92,9	72,7	414	325	78,5	325	287	88,3	69,3	71,1	
	5942	316	258	81,6	258	248	96,1	78,5	309	255	82,5	255	231	90,6	74,8	76,6	
	5943	144	113	78,5	113	111	98,2	77,1	165	138	83,6	138	126	91,3	76,4	76,7	
	5951	252	205	81,3	205	185	90,2	73,4	233	185	79,4	185	168	90,8	72,1	72,8	
	5952	262	183	69,8	183	167	91,3	63,7	280	227	81,1	227	200	88,1	71,4	67,7	
	5953	234	179	76,5	179	141	78,8	60,3	160	129	80,6	129	113	87,6	70,6	64,5	
Yn	6001	577	522	90,5	522	485	92,9	84,1	128	110	85,9	110	98	89,1	76,6	82,7	

		Base aréolaire							Bases téléphoniques							
Prov. Terr.	Région socio-sanitaire	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné
T.N.-O.	6101	604	509	84,3	509	470	92,3	77,8	90	70	77,8	70	69	98,6	76,7	77,7
Nt	6201	386	349	90,4	349	327	93,7	84,7	84,7

Taux de réponse par réseau local d'intégration des services de santé (RLISS) et par base de sondage de l'ESCC 2009 en Ontario

		Base aréolaire							Bases téléphoniques							
Prov. Terr.	RLISS	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	Taux de rép. combiné
Ont.	Total	13 662	11 229	82,2	11 229	10 211	90,9	74,7	15 703	12 256	78,0	12 256	10 758	87,8	68,5	71,4
	3501	885	759	85,8	759	711	93,7	80,3	1 076	854	79,4	854	747	87,5	69,4	74,3
	3502	1 611	1 381	85,7	1 381	1 309	94,8	81,3	1 761	1 405	79,8	1 405	1 253	89,2	71,2	76,0
	3503	779	666	85,5	666	609	91,4	78,2	826	660	79,9	660	593	89,8	71,8	74,9
	3504	1 651	1 322	80,1	1 322	1 180	89,3	71,5	1 798	1 389	77,3	1 389	1 224	88,1	68,1	69,7
	3505	726	602	82,9	602	531	88,2	73,1	823	629	76,4	629	528	83,9	64,2	68,4
	3506	749	636	84,9	636	583	91,7	77,8	843	653	77,5	653	570	87,3	67,6	72,4
	3507	752	574	76,3	574	533	92,9	70,9	873	636	72,9	636	546	85,8	62,5	66,4
	3508	859	708	82,4	708	634	89,5	73,8	1 136	860	75,7	860	732	85,1	64,4	68,5
	3509	1 217	1 005	82,6	1 005	899	89,5	73,9	1 483	1 160	78,2	1 160	996	85,9	67,2	70,2
	3510	789	651	82,5	651	579	88,9	73,4	935	743	79,5	743	652	87,8	69,7	71,4
	3511	1 289	1 011	78,4	1 011	921	91,1	71,5	1 364	1 076	78,9	1 076	956	88,8	70,1	70,8
	3512	594	460	77,4	460	420	91,3	70,7	732	571	78,0	571	512	89,7	69,9	70,3
	3513	1 111	956	86,0	956	845	88,4	76,1	1 325	1 037	78,3	1 037	926	89,3	69,9	72,7

	3514	650	498	76,6	498	457	91,8	70,3	728	583	80,1	583	523	89,7	71,8	71,1
--	------	-----	-----	------	-----	-----	------	------	-----	-----	------	-----	-----	------	------	------

Taux de réponse par province et par base de sondage pour le sous-échantillon de l'enquête sur l'accès aux services de soins de santé (EASS) 2009

Prov. Terr.	Base aéroilaire							Bases téléphoniques							Taux de rép. combiné
	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	# mén. cibles	# mén. rép.	Taux de rép. mén.	# pers. sélect.	# rép.	Taux de rép. pers.	Taux de rép.	
Canada	30 394	25 145	82,7	23 732	21 648	91,2	75,5	33 459	26 713	79,8	24 870	21 963	88,3	70,5	72,9
T.-N.-L.	1 141	991	86,9	935	855	91,4	79,4	1 336	1 137	85,1	1 033	899	87,0	74,1	76,6
Î.-P.-É.	620	524	84,5	480	430	89,6	75,7	645	545	84,5	501	427	85,2	72,0	73,8
N.-É.	1 531	1 308	85,4	1 223	1 106	90,4	77,3	1 560	1 338	85,8	1 244	1 105	88,8	76,2	76,7
N.-B.	1 493	1 283	85,9	1 215	1 090	89,7	77,1	1 605	1 387	86,4	1 300	1 169	89,9	77,7	77,4
Qc	3 006	2 374	79,0	2 234	2 091	93,6	73,9	3 481	2 727	78,3	2 578	2 265	87,9	68,8	71,2
Ont.	13 662	11 229	82,2	10 649	9 676	90,9	74,7	15 051	11 747	78,0	10 903	9 548	87,6	68,3	71,4
Man.	1 996	1 671	83,7	1 579	1 430	90,6	75,8	1 997	1 676	83,9	1 558	1 421	91,2	76,5	76,2
Sask.	1 905	1 696	89,0	1 592	1 507	94,7	84,3	2 348	1 926	82,0	1 802	1 650	91,6	75,1	79,2
Alb.	2 364	1 906	80,6	1 809	1 606	88,8	71,6	2 590	2 059	79,5	1 922	1 708	88,9	70,6	71,1
C.-B.	2 676	2 163	80,8	2 016	1 857	92,1	74,5	2 846	2 171	76,3	2 029	1 771	87,3	66,6	70,4