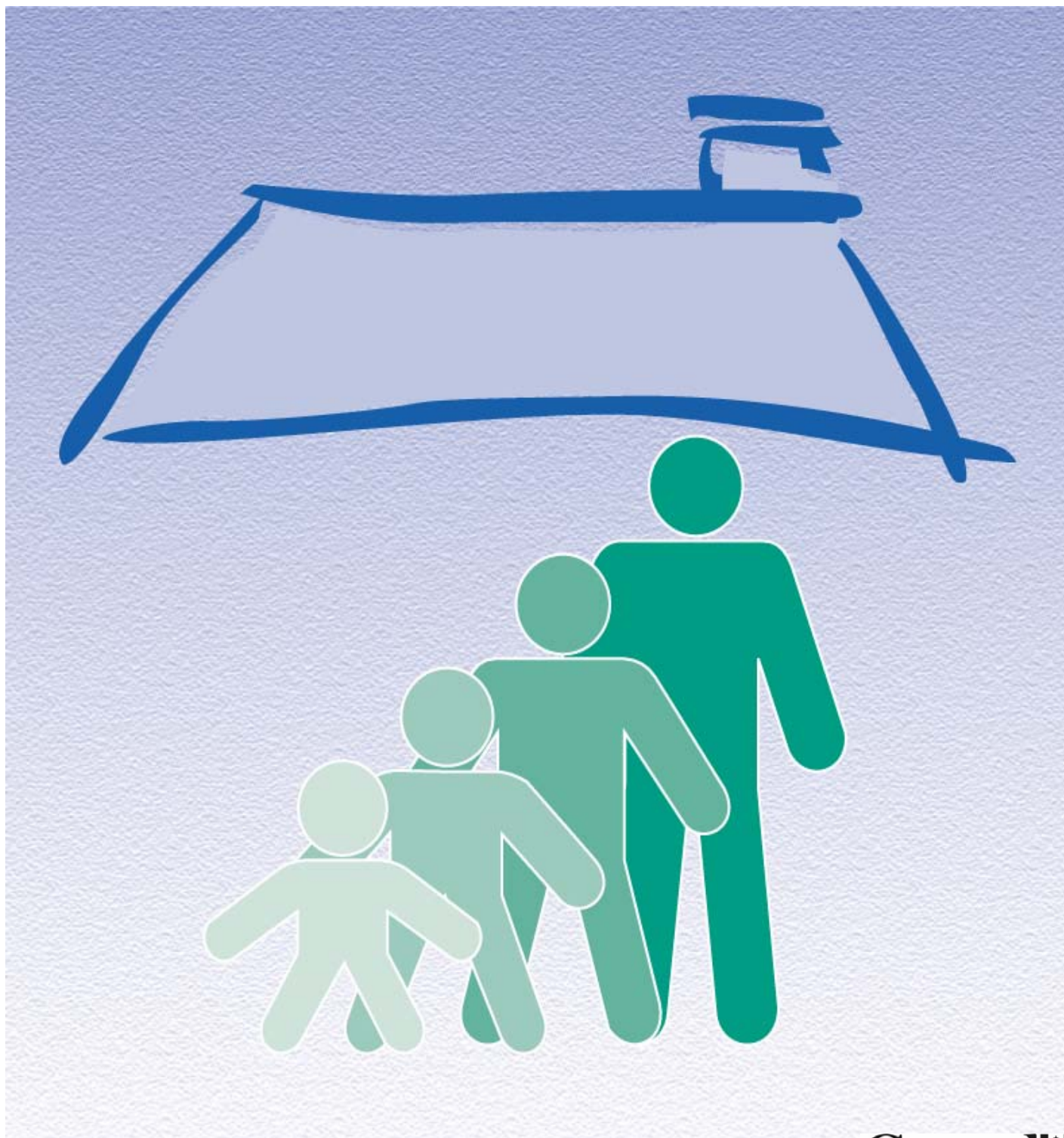


Enquête sur la santé dans les collectivités canadiennes (ESCC)

Fichiers de microdonnées 2007
Guide de l'utilisateur

Juin 2008



Statistique
Canada

Statistics
Canada

Canada

TABLE DES MATIÈRES

1.	INTRODUCTION	1
2.	HISTORIQUE	2
3.	REMANIEMENT DE L'ESCC EN 2007	3
4.	STRUCTURE DU CONTENU DE L'ESCC	4
4.1	CONTENU COMMUN.....	4
4.2	CONTENU OPTIONNEL	5
4.3	CONTENU DE RÉPONSE RAPIDE.....	5
4.4	CONTENU DES FICHIERS DE MICRODONNÉES	5
5.	PLAN D'ÉCHANTILLONNAGE	8
5.1	POPULATION CIBLE	8
5.2	RÉGIONS SOCIO-SANITAIRES.....	8
5.3	TAILLE ET RÉPARTITION DE L'ÉCHANTILLON	8
5.4	BASES DE SONDAGE ET STRATÉGIES D'ÉCHANTILLONNAGE DES MÉNAGES	10
5.4.1	ÉCHANTILLONNAGE DES MÉNAGES À PARTIR DE LA BASE ARÉOLAIRE	10
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	13
5.5	RÉPARTITION DE L'ÉCHANTILLON PAR PÉRIODE DE COLLECTE DES DONNÉES	14
5.6	ÉCHANTILLONNAGE DES PERSONNES INTERVIEWÉES	14
5.7	ACHAT D'ÉCHANTILLON SUPPLÉMENTAIRE POUR 3 RÉGIONS SOCIO-SANITAIRES EN ONTARIO.....	15
5.8	SOUS-ÉCHANTILLONNAGE POUR L'ENQUÊTE SUR L'ACCÈS AUX SERVICES DE SANTÉ (EASS).....	16
6.	COLLECTE DES DONNÉES	19
6.1	INTERVIEWS ASSISTÉES PAR ORDINATEUR.....	19
6.2	DÉVELOPPEMENT DES APPLICATIONS DE L'ESCC.....	19
6.3	FORMATION DES INTERVIEWEURS	20
6.4	L'INTERVIEW	21
6.5	OPÉRATIONS SUR LE TERRAIN	24
6.6	CONTRÔLE DE LA QUALITÉ ET GESTION DE LA COLLECTE	24
7.	TRAITEMENT DES DONNÉES	26
7.1	VÉRIFICATION.....	26
7.2	CODAGE.....	26
7.3	CRÉATION DE VARIABLES DÉRIVÉES	26
7.4	PONDÉRATION.....	27
8.	PONDÉRATION	28
8.1	INTRODUCTION	28
8.2	PONDÉRATION DE L'ÉCHANTILLON PROVENANT DE LA BASE ARÉOLAIRE	29
8.3	PONDÉRATION DE L'ÉCHANTILLON PROVENANT DE LA BASE TÉLÉPHONIQUE	31
8.4	INTÉGRATION DES BASES ARÉOLAIRE ET TÉLÉPHONIQUE	33
8.5	ÉTAPES DE PONDÉRATION POST-INTÉGRATION.....	33
8.6	PARTICULARITÉS DE LA PONDÉRATION POUR LES TROIS TERRITOIRES	35
9.	QUALITÉ DES DONNÉES	37
9.1	TAUX DE RÉPONSE	37
9.2	ERREURS DANS LES ENQUÊTES.....	38
9.2.1	ERREURS NON DUES À L'ÉCHANTILLONNAGE.....	38
9.2.2	ERREURS DUES À L'ÉCHANTILLONNAGE	39

10.	LIGNES DIRECTRICES POUR LA TOTALISATION, L'ANALYSE ET LA DIFFUSION	40
10.1	LIGNES DIRECTRICES POUR L'ARRONDISSEMENT	40
10.2	LIGNES DIRECTRICES POUR LA PONDÉRATION DE L'ÉCHANTILLON EN VUE DE LA TOTALISATION	41
10.2.1	DÉFINITIONS DES CATÉGORIES D'ESTIMATIONS : DE TYPE NOMINAL PAR OPPOSITION À QUANTITATIVES.....	41
10.2.2	TOTALISATION D'ESTIMATIONS DE TYPE NOMINAL.....	42
10.2.3	TOTALISATION D'ESTIMATIONS QUANTITATIVES	42
10.3	LIGNES DIRECTRICES POUR L'ANALYSE STATISTIQUE	43
10.4	LIGNES DIRECTRICES POUR LA DIFFUSION	44
11.	TABLEAUX DE LA VARIABILITÉ D'ÉCHANTILLONNAGE APPROXIMATIVE	46
11.1	COMMENT UTILISER LES TABLEAUX DE CV POUR LES ESTIMATIONS DE TYPE NOMINAL	46
11.2	EXEMPLES D'UTILISATION DES TABLEAUX DE CV POUR DES ESTIMATIONS DE TYPE NOMINAL.....	49
11.3	COMMENT UTILISER LES TABLEAUX DE CV POUR CALCULER LES LIMITES DE CONFIANCE.....	52
11.4	EXEMPLE D'UTILISATION DE TABLEAUX DE CV POUR OBTENIR DES LIMITES DE CONFIANCE	53
11.5	COMMENT UTILISER LES TABLEAUX DE CV POUR EFFECTUER UN TEST Z	53
11.6	EXEMPLE D'UTILISATION DES TABLEAUX DE CV POUR EFFECTUER UN TEST Z	54
11.7	VARIANCES OU COEFFICIENTS DE VARIATION EXACTS	54
11.8	SEUILS POUR LA DIFFUSION DES ESTIMATIONS RELATIVES À L'ESCC.....	55
12.	FICHIERS DE MICRODONNÉES : DESCRIPTION, ACCÈS ET UTILISATION.....	56
12.1	FICHIERS MAÎTRES	56
12.1.1	CENTRES DE DONNÉES DE RECHERCHE.....	56
12.1.2	TÉLÉ-ACCÈS.....	56
12.1.3	TOTALISATIONS PERSONNALISÉES	56
12.2	FICHIERS DE PARTAGE	57
12.3	FICHIERS DE MICRODONNÉES À GRANDE DIFFUSION	57
12.4	UTILISATION DE LA VARIABLE DE PONDÉRATION.....	57
12.5	CONVENTION APPLIQUÉE POUR NOMMER LES VARIABLES À PARTIR DE 2007.....	58
12.6	CONVENTION APPLIQUÉE POUR NOMMER LES VARIABLES AVANT 2007	60
12.7	LIGNES DIRECTRICES POUR L'UTILISATION DES VARIABLES D'UN SOUS-ÉCHANTILLON	61
12.7.1	CONVENTION DE DÉNOMINATION DES VARIABLES DES MODULES DE SOUS-ÉCHANTILLON.....	61
12.7.2	FICHIER DISTINCT	62
12.7.3	ANALYSE COMBINANT LE CONTENU DES SOUS-ÉCHANTILLONS ET LE CONTENU OPTIONNEL	63
12.8	DICIONNAIRES DE DONNÉES	63
12.9	DIFFÉRENCES DANS LE CALCUL DES VARIABLES DE CONTENU COMMUN FONDÉ SUR DIFFÉRENTS FICHIERS.....	63
ANNEXE A	CONTENU DE L'ESCC POUR 2007 ET 2008	65
ANNEXE B	SÉLECTION DU CONTENU OPTIONNEL SELON LA PROVINCE ET LE TERRITOIRE, 2007	69
ANNEXE C	GÉOGRAPHIE DE L'ESCC	75
ANNEXE D	RÉPARTITION DE L'ÉCHANTILLON PAR RÉGION	81
ANNEXE E	TAUX DE RÉPONSE PAR RÉGION SOCIO SANITAIRE ET PAR BASE DE SONDAJE	87

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 qui vivent dans les réserves et dans d'autres peuplements autochtones des provinces, les pensionnaires d'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. Il importe toutefois de mentionner qu'un FMGD n'a pas été produit pour l'ESCC 2007. Le FMGD sera produit à l'été 2009 lorsque les données de 2007 et 2008 seront disponibles.

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.ca
Renseignements sur le télé-accès :	613-951-1746
Courriel :	cchs-escc@statcan.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 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 tout simplement par le nom de l'enquête suivi de la période de collecte (par exemple, « Enquête sur la santé dans les collectivités canadiennes, 2007 » ou « ESCC – 2007 »). 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 deux : le contenu de base et le contenu thématique. La composante thématique a été créée afin de tirer profit de l'approche de collecte continue. La durée de collecte des données de cette composante pourra être adaptée selon la prévalence des estimations désirées et leur niveau géographique souhaité. Le contenu de base demeurera 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 qu'à chaque 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. À compter de 2008, les données de l'ESCC seront diffusées sur une base annuelle. À tous les deux ans, un fichier regroupant l'échantillon des deux années (taille de 130 000) sera 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. **STRUCTURE DU 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 de base et le contenu thématique, 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 2007 selon leur composante.

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

Tableau 4.1 Durée de l'enquête selon les composantes

Composante de l'ESCC	Durée moyenne
Contenu commun <ul style="list-style-type: none">• Contenu de base• Contenu thématique	30 minutes (20 minutes) (10 minutes)
Contenu optionnel	10 minutes
Contenu de réponse rapide (facultatif)	2 minutes

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). Elle est divisée en deux volets : le contenu de base et le contenu thématique.

Le contenu de base est formé de questions qui sont demandées à tous les répondants de l'enquête. Ces questions demeureront relativement inchangées dans le questionnaire pour une période d'environ 6 ans, à moins qu'un souci important vis-à-vis la qualité le requiert.

Le contenu thématique est formé de questions qui sont reliées à un sujet spécifique. Le contenu thématique est formé d'un thème de deux ans et d'un thème annuel et compte pour 10 minutes de l'entrevue. Certains thèmes pourraient être réintroduits dans l'enquête à tous les 2, 4 ou 6 ans au besoin. Cette composante permet à l'ESCC une meilleure planification du contenu à moyen terme.

Certains modules inclus dans la composante thématique 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 le cycle 3.1 de l'ESCC, 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 2007 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

Les fichiers de microdonnées collectés en 2007 sont disponibles depuis juin 2008. En juin 2009, deux fichiers principaux seront disponibles : un fichier principal de données fondé sur la période de collecte de 2008, qui s'apparentera au fichier de données de 2007, ainsi qu'un fichier de données fondé sur la période de collecte de 2007-2008. Le fichier principal de 2007-2008 sera de la même taille que les fichiers des cycles précédents (environ 130 000 répondants). Il comprendra le contenu de base, le contenu optionnel et le contenu thématique recueillis au cours de la période de deux ans. Les thèmes touchant une année ne seront pas disponibles dans le fichier de données de deux ans. Par ailleurs, les modules thématiques collectés auprès d'un sous-échantillon de la population continueront d'être diffusés dans des fichiers séparés. Ces derniers incluent le contenu de base et le contenu des modules thématiques collectés auprès d'un sous-échantillon de

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.

répondants. Le tableau 4.2 fournit des précisions concernant ce qui sera disponible avec les diffusions de 2007 et 2008.

Tableau 4.2 Composantes de contenu comprises dans les fichiers de données de 2007 et 2008

Fichiers		Contenu de base	Contenu thématique de 2007 ¹	Contenu thématique de 2008 ²	Contenu thématique de 2007-2008	Contenu optionnel ³
Fichiers de 2007	Principal	Oui	Non	S/O	Oui	Oui
	Sous-échantillon (3 modules)	Oui	Oui	S/O	Non	Non
Fichiers de 2008	Principal	Oui	S/O	Oui (sauf le module de sous-échantillon)	Oui	Oui
	Sous-échantillon (1 module)	Oui	S/O	Un module (Poids et taille mesurés)	Non	Non
Fichier de 2007-2008	Principal	Oui	Non	Non	Oui	Oui

1. Le thème de 2007 était composé de 3 modules (Satisfaction des patients, Accès aux services de santé et Temps d'attentes) qui ont été posés à un sous-échantillon de répondants.

2. Le thème de 2008 comprend un groupe de modules reliés à la détection des maladies chroniques ainsi qu'un module sur le poids et la taille mesurés. Ce dernier est demandé à un sous-échantillon de répondants.

3. Selon le principe que le contenu optionnel demeurera le même pendant deux ans. Autrement, il sera inclus uniquement dans le fichier pour l'année pendant laquelle il a été recueilli.

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 qui vivent dans les réserves et dans d'autres peuplements autochtones des provinces, les pensionnaires d'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 (RSS) et chaque territoire est considéré comme formant une RSS unique. En collaboration avec les provinces, Statistique Canada est parfois appelé à modifier les limites de certaines RSS 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 2007, des données ont été recueillies pour 118 RSS dans les 10 provinces, ainsi que pour une RSS par territoire, soit, en tout, pour 121 RSS (Annexe C).

5.3 Taille et répartition de l'échantillon

Afin de produire des estimations fiables pour chaque RSS et compte tenu du budget accordé pour l'ESCC, il a été établi que cette composante de l'enquête doit être réalisée auprès d'un échantillon d'environ 130 000 personnes sur une période de 2 ans. La production d'estimations fiables pour chaque RSS 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 RSS et aux provinces. A la première étape, on impose une taille minimum de 500 répondants par RSS. C'est le minimum considéré 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 RSS 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. L'échantillon est ensuite divisé également pour chacune des 2 années de collecte. Le tableau 5.1 donne la taille d'échantillon annuel pour 2007.

Tableau 5.1 Nombre de régions sociosanitaires et tailles visées d'échantillon selon la province/territoire, 2007

Province	Nombre de RSS	Taille totale de l'échantillon (visée)
Terre-Neuve-et-Labrador	4	2 005
Île-du-Prince-Édouard ¹	3	1 480
Nouvelle-Écosse	6	2 520
Nouveau-Brunswick	7	2 580
Québec	16	12 145
Ontario ²	36	22 230
Manitoba	10	3 755
Saskatchewan	11	3 860
Alberta	9	6 105
Colombie-Britannique	16	8 050
Yukon	1	600
Territoires du Nord-Ouest	1	600
Nunavut	1	350
Canada	121	66 280

1. La taille de l'échantillon pour l'Île-du-Prince-Édouard a été augmentée pour répondre aux besoins de l'Enquête sur l'accès aux services de soins de santé (EASS) de 2007. La taille initiale était de 1 000 unités (se référer à la section 5.8 pour plus de détails).
2. La 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 20 900 unités (se référer à la section 5.7 pour plus de détails)

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

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 2007, 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 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 regroupent un ensemble de collectivités (pour plus de détails, se reporter à la section 5.4.1).

L'échantillon est ensuite divisé également en 2 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 RSS⁴, tel qu'élaboré dans la section suivante. Mentionnons finalement que l'effectif des échantillons tirés de chaque base a été augmenté avant la collecte des données afin de tenir compte des unités hors du champ de l'enquête et du taux prévu de non-réponse à partir des taux obtenus des cycles précédents de l'ESCC. Les tailles d'échantillons par RSS 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 2007 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 des bâtiments d'habitation (parfois 12 ou 18 appartements) sont sélectionnées 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.

4. À 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).

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

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 aires de diffusion 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'unité primaire d'échantillonnage 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 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, 2007

Région	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 2007, il a fallu sélectionner plus de 47 000 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 2007, il fallait obtenir un échantillon d'environ 47 000 logements réparti par RSS dont les limites géographiques différaient de celles des régions économiques de l'EPA. Globalement, l'ESCC 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 0,78 (47 000/60 000). Toutefois, comme ce facteur de redressement variait de 0,3 à 3,0 au niveau des RSS, certains ajustements ont été nécessaires.

Les modifications apportées au processus de sélection dans les régions variaient selon le facteur de redressement. Pour les RSS où le facteur était égal ou inférieur à 1, le nombre d'UPE choisi 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 RSS 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 RSS 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 RSS. 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 RSS en question⁶.

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 a consisté 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 plus 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 2007 couvrait 90 % des ménages privés du Yukon, 97 % de ceux des Territoires du Nord-Ouest et 71 % de ceux du Nunavut⁷.

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.

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.

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

À l'exception de 5 RSS (les 2 RSS 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 téléphones 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 RSS en utilisant un fichier de conversion de codes postaux de sorte à pouvoir associer un RSS à chaque numéro de téléphone. Dans chaque RSS, 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 la sous-couverture de la base liste de numéros de téléphone est plus importante que celle 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 de la sous-couverture dû à l'utilisation de la base liste de numéros de téléphone est minimal et est corrigé lors de la pondération.

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

Dans 4 RSS, 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.

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.

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 RSS 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 RSS, 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 RSS 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 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. Comme pour le cycle 3.1, l'ESCC 2007 a continué de sélectionner une personne 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 (deux adultes âgés entre 45 et 64 et un jeune de 15 ans), le jeune de 15 ans aurait 5 fois plus de chances d'être sélectionné que les 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

Facteurs multiplicatifs du poids de sélection					
Age	12-19	20-29	30-44	45-64	65+
Facteur	5	2	2	1	1

5.7 Achat d'échantillon supplémentaire pour 3 régions sociosanitaires en 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 (voir Annexe C). 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 RSS et RLISS s'entrecoupent, le niveau de stratification utilisé a été le croisement RSS-RLISS. Les tailles d'échantillon préalablement alloué par RSS ont donc été conservées. Dans les cas où la répartition par RSS 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 RSS-RLISS. Le tableau 5.4 donne la taille d'échantillon de répondants visés par RLISS pour l'année 2007.

Tableau 5.4 Tailles visées de répondants par réseau local d'intégration des services de santé (RLISS), ESCC 2007

RLISS	Taille totale de l'échantillon (visée)
01-Erie St. Clair	1 552
02-South West	2 568
03-Waterloo Wellington	1 230
04-Hamilton Niagara Haldimand Brant	2 620
05-Central West	1 062
06-Mississauga Halton	1 136
07-Toronto Central	1 084
08-Central	1 404
09-Central East	2 076
10-South East	1 344
11-Champlain	2 058
12-North Simcoe Muskoka	1 058
13-North East	1 992
14-North West	1 044
Ontario	22 228

La taille totale de l'échantillon des croisements RSS-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 34 000 répondants, permettant de produire des estimations fiables au niveau provincial. De plus, les mêmes tailles d'échantillon étaient visées par province. Par contre, pour les provinces Maritimes, la taille d'échantillon disponible de l'ESCC était inférieure à la taille de l'échantillon de l'EASS 2005. En effet, l'ESCC 2007 ne couvre que la moitié de l'échantillon des cycles précédents étant donné sa stratégie de collecte continue. A l'exception de l'Île-du-Prince-Édouard (I-P-E), les autres provinces maritimes n'ont utilisé que la taille disponible de l'ESCC 2007. Dans la province de l'I-P-E, l'échantillon de l'ESCC 2007 a été augmenté de 1 000 à 1 480 répondants afin de répondre aux besoins de l'EASS. De plus, l'Ontario a décidé d'utiliser l'échantillon complet de l'ESCC pour l'EASS. Voici donc les tailles finales pour le sous-échantillon de l'EASS sélectionné à partir de l'ESCC 2007 et la comparaison avec les tailles de l'EASS 2005.

Table 5.5 Tailles du sous-échantillon sélectionné par province pour l'EASS à partir de l'ESCC 2007 et comparaison avec les tailles de l'EASS 2005.

Province	Taille d'échantillon		
	ESCC 2007	EASS 2007	EASS 2005
Terre-Neuve-et-Labrador	2 005	2 005	2 700
Île-du-Prince-Édouard	1 480	1 480	1 850
Nouvelle-Écosse	2 520	2 520	3 000
Nouveau-Brunswick	2 580	2 580	3 000
Québec	12 145	4 600	4 600
Ontario	22 230	22 230	4 800
Manitoba	3 755	3 200	3 200
Saskatchewan	3 860	3 200	3 200
Alberta	6 105	3 600	3 600
Colombie-Britannique	8 050	4 000	4 000
CANADA	66 280	49 415	33 950

Une fois la taille définie par province, l'échantillon a été réparti par RSS proportionnellement à la taille de la population par RSS, permettant ainsi une meilleure répartition de l'échantillon par province tout en tenant compte de la stratification de l'ESCC par RSS. Pour les provinces où la taille d'échantillon par RSS était insuffisante, une répartition de puissance avec puissance inférieure à 1 a dû être utilisée. Une puissance de 0,9 a été utilisée au Québec, 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 RSS 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 de 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 2007 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 2007. Un sous-échantillon de logements ou de numéros de téléphone de même taille a été tiré aléatoirement dans chaque RSS à chaque période de collecte.

6. COLLECTE DES DONNÉES

6.1 Interviews assistées par ordinateur

Entre janvier et décembre 2007, un total de 65 946 interviews valables assistées par ordinateur (IAO) ont été effectuées. Environ la moitié des interviews 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

Avant le début de la collecte, des gestionnaires de projet des bureaux régionaux de collecte ont assisté à des séances de formation à Statistique Canada. Les séances, animées par des membres de l'équipe responsable du projet de l'ESCC, ont consisté en des descriptions de la formation qui serait donnée aux intervieweurs en région. Les gestionnaires ont ensuite animé des séances de formation personnalisées à l'intention des intervieweurs. Des membres de l'équipe responsable de l'enquête

ont assisté aux séances afin d'y offrir un soutien et de répondre aux questions et aux préoccupations des participants.

L'objet des séances était de familiariser les intervieweurs avec les applications de l'ESCC 2007, notamment le contenu. 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, près des trois quarts (72,1 %) 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. De plus, les questions ont été traduites en chinois, panjabi et inuktitut. Le chinois et le panjabi sont les deux langues qui ont le plus souvent posé problème pour les bureaux régionaux. La traduction en inuktitut a servi à la collecte de données au Nunavut.

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 2007 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 78 % a été atteint. L'annexe D 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 2007. Il s'agissait, 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 IPAQ, 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'IAQ afin de recenser les interviews de durée excessivement courte. Ces dernières ont été signalées aux moyens d'indicateurs, supprimées des microdonnées et classées parmi les cas de non-réponse.

Par ailleurs, des rapports personnalisés ont été établis et utilisés pour examiner des questions particulières de qualité des données qui se sont manifestées au cours de la collecte. À titre d'exemple, le protocole de l'ESCC permet des interviews par procuration lorsque le répondant choisi est mentalement ou physiquement incapable de mener l'interview. Les interviews par procuration ont une valeur limitée étant donné que plusieurs modules ne sont pas demandés pour des raisons reliées à la qualité des données. Il est donc important de déterminer les taux d'interview par procuration et les motifs derrière ces interviews. Un rapport d'interview par procuration a été créé afin de suivre leur progression. Ces rapports ont permis d'identifier les intervieweurs qui semblaient mal comprendre les circonstances qui rendaient la conduite d'une interview par procuration acceptable. Ces derniers ont bénéficié ensuite d'un rappel ou d'une formation supplémentaire axée sur les critères particuliers à appliquer dans le cas d'interview par procuration.

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 norme 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 2007. 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 du cycle 3.1 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 les fichiers de microdonnées et doit servir à extraire des estimations de l'enquête. Par exemple, si l'on doit évaluer le nombre de personnes qui fument tous les jours, on le fait en choisissant dans l'échantillon les enregistrements des personnes qui présentent cette caractéristique et en faisant la somme des facteurs de pondération que représentent ces enregistrements.

Vous trouverez les détails sur la façon dont on calcule les poids d'échantillonnage à la section 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.

La stratégie de pondération a été développée en traitant séparément la base aréolaire et la base téléphonique. Les poids des ménages résultant de ces deux bases sont ensuite combinés en un seul ensemble de poids des ménages lors d'une étape appelée « *intégration* ». Suite à la transformation des poids des ménages en poids de personne et à quelques autres ajustements, ce poids intégré devient le poids de personne final. Noter que, selon le besoin, une seule ou deux bases peuvent être utilisées pour la sélection de l'échantillon dans une région sociosanitaire (RSS) donnée. La stratégie de pondération tient compte de cette particularité lors de l'étape d'intégration.

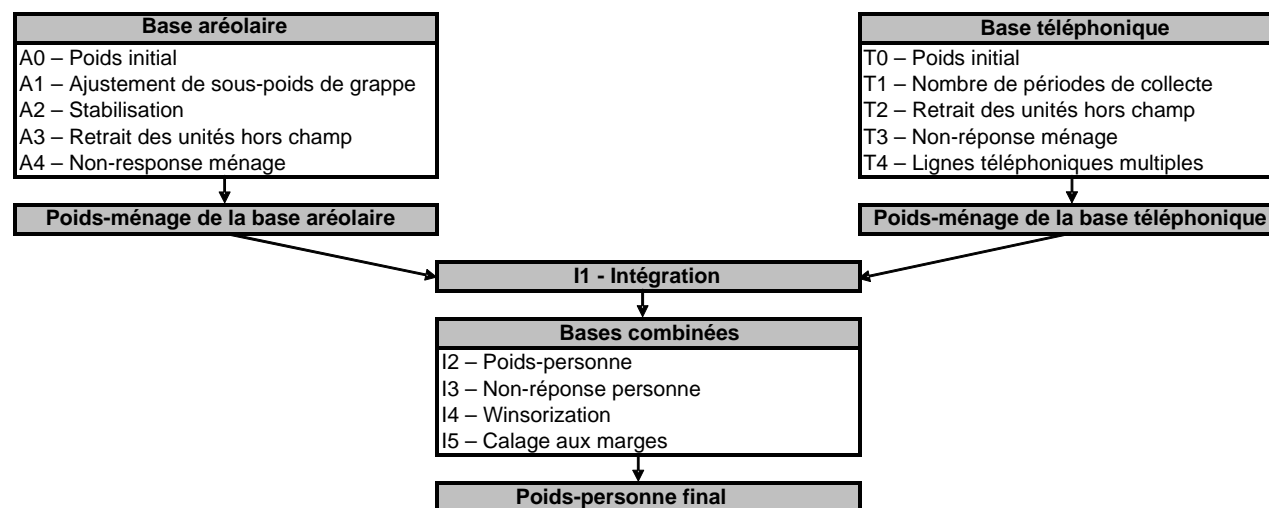
8.1 Introduction

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, suit l'étape finale de la pondération, c'est-à-dire le calage aux marges où les poids sont ajustés afin de correspondre à des comptes de population connus et, dans un même temps, de contrôler la saisonnalité des données. Cette étape est expliquée 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'EPA. Ce poids est basé sur le plan de sondage de l'EPA puisque l'échantillon 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, 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 (RSS). 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 A0. Pour plus de détails sur le mécanisme de sélection, de même qu'une définition plus complète des strates et des grappes, se référer à Statistique Canada 1998⁹.

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 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 l'ESCC. 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).

9. 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 RSS, l'accroissement de la taille 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 RSS 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 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'intérieur de groupes homogènes de réponse (GHR). Dans le but de créer ces GHR, 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 par groupes ayant des propriétés de réponse similaires. 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 ainsi que des para-données ce qui inclut 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. Un facteur d'ajustement a donc été calculé à l'intérieur de chaque GHR 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 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. Noter qu'une seule de ces deux bases peut être utilisée à l'intérieur d'une RSS. La base liste est toujours utilisée comme complément à la base aréolaire tandis que la base CA est toujours utilisée seule pour une RSS 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 RSS. C'est pourquoi certaines unités ont été échantillonnées dans une certaine RSS alors que l'information fournie lors de l'interview les localisait plutôt dans une RSS avoisinante. Cette particularité a été contournée lors de la pondération en appliquant les premiers trois ajustements (T0, T1 et T2) relativement à la RSS assignée lors de la sélection de l'échantillon. Les 2 ajustements restants (T3 et T4) sont appliqués en se basant sur la RSS provenant de l'information recueillie du répondant pour s'assurer que toutes les unités font partie de leur vraie RSS.

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 Paton¹⁰ 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 RSS. 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 RSS. 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.

10. 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 correspond un poids initial faisant en sorte que chaque échantillon soit représentatif de la RSS. 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. Le facteur d'ajustement appliqué à chaque échantillon est égal à l'inverse du nombre d'échantillons combinés ou encore, 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 afin de 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 classe a été obtenu de la façon suivante:

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

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'intégration¹¹. 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, compris entre 0 et 1, est déterminé au niveau de RSS de façon à représenter l'importance relative de chaque échantillon dans l'échantillon total. Cette importance relative est mesurée en termes de taille d'échantillon. Plus la proportion d'échantillon qu'une base représente dans l'échantillon total est grande, plus grande sera son importance relative dans l'échantillon total. Pour obtenir le facteur d'ajustement d'intégration, on calcule d'abord un facteur α , obtenu de la façon suivante:

$$\alpha = n_A / (n_A + n_T)$$

où n_A et n_T représentent respectivement les tailles d'échantillon des bases aréolaire et téléphonique. 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 RSS n'est couverte que par une seule base, le facteur d'ajustement est égal à 1. 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 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 5.5 pour plus de détails).

11. 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.

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 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 classes définies à partir des caractéristiques disponibles pour les répondants et non-répondants. Toutes les caractéristiques recueillies lors du listage des membres du ménage, en plus de l'information géographique et des para-données, étaient en fait disponibles pour créer ces classes. La méthode du score est utilisée afin de définir les classes. Un facteur d'ajustement est calculé à l'intérieur de chaque classe 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 a donc été 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

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

15- Calage aux marges

La dernière étape nécessaire afin d'obtenir le poids final de l'ESCC est le calage aux marges (I5). Le calage est appliqué en utilisant CALMAR¹² afin de s'assurer que la somme des poids finaux corresponde aux estimations de populations définies à l'échelle des RSS, 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. Un minimum de 20 répondants est nécessaire dans chaque domaine pour le calage au niveau RSS, par groupe âge-sexe. Pour les domaines ayant moins de 20 répondants, des regroupements sont effectués au niveau des provinces et/ou du sexe. Dans un même temps, les poids sont ajustés afin de s'assurer que chaque période de collecte (de deux mois) est également représentée dans 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 comptes du Recensement les plus récents, de même que sur les comptes de naissance, décès, immigration et émigration depuis ce temps. La moyenne des estimations mensuelles pour chacun des croisements RSS-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 final de l'ESCC* que l'on retrouve dans le fichier de données portant le nom de variable WTS_M.

8.6 Particularités de la pondération pour les trois territoires

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. Noter que les capitales de chaque territoire formaient une strate à elles seules, et sont donc toutes trois sélectionnées automatiquement à 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.

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

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 un calage aux marges est appliqué de façon semblable à ce qui est fait pour les provinces, à l'exception de deux 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 la sous couverture de la base aréolaire, telle que décrite à la section 5.4.1.

9. QUALITÉ DES DONNÉES

9.1 Taux de réponse

Au total, 84 973 des unités sélectionnées à l'ESCC 2007 faisaient partie du champ de l'enquête¹³. Parmi ces unités, 71 922 ont accepté de participer à l'enquête ce qui résulte en un taux de réponse à l'échelle du ménage de 84,6%. Parmi ces ménages répondants, 71 922 personnes ont été sélectionnées (une personne par ménage) pour participer à l'enquête parmi lesquelles 65 946 ont accepté de le faire ce qui résulte en un taux de réponse à l'échelle de la personne de 91,7 %. À l'échelle canadienne, un taux de réponse combiné de 77,6 % a donc été observé pour l'ESCC 2007. L'Annexe E 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 de même que par réseau local d'intégration des services de santé (RLISS) en Ontario et par province pour le sous-échantillon de l'Enquête sur l'accès aux services de santé (EASS).

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.

Taux de réponse à l'échelle du ménage

HHRR =	<i># de ménages répondants provenant des 2 bases</i>
	<i>tous les ménages faisant partie du champ de l'enquête provenant des 2 bases</i>

Taux de réponse à l'échelle de la personne

PPRR =	<i># de répondants provenant des 2 bases</i>
	<i>toutes les personnes sélectionnées provenant des 2 bases</i>

Taux de réponse combiné = HHRR x PPRR

13. Parmi les unités sélectionnées au départ, certaines ne font pas parti du champ de l'enquête. Ce sont par exemples des logements vacants ou démolis, des logements non-résidentiels ou encore des numéros de téléphones non-valides 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.

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

$$\text{HHRR} = \frac{35\,748 + 36\,174}{40\,668 + 44\,310} = \frac{71\,922}{84\,978} = 0,846$$

$$\text{PPRR} = \frac{33\,473 + 32\,473}{35\,748 + 36\,174} = \frac{65\,946}{71\,922} = 0,917$$

$$\begin{aligned} \text{Taux de réponse combiné} &= 0,846 \times 0,917 \\ &= 0,776 \\ &= \mathbf{77,6\%} \end{aligned}$$

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 non-réponse partielle (le fait de ne pas répondre à une ou plusieurs questions) à une non-réponse totale. Dans l'ESCC 2007, 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 socio-sanitaires, l'échantillon initial a été augmenté au cours de l'année afin d'atteindre les tailles ciblées pour 2007. 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 2007 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-section 10.4.

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 (Maître, Partagé, 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 finals 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 finals des enregistrements contenant la caractéristique voulue pour le numérateur (\hat{X});
- b) additionner les poids finals 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 final, puis faire la somme de cette quantité pour tous les enregistrements visés pour obtenir le numérateur (\hat{X});
- b) faire la somme des poids finals 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.

Pour permettre à l'utilisateur d'évaluer la qualité des totalisations estimées d'après les données, Statistique Canada a produit un ensemble de tableaux de variabilité d'échantillonnage approximative (couramment appelées « Tableaux des CV ») pour l'ESCC. On peut employer ces tableaux pour obtenir des coefficients de variation approximatifs pour les estimations de type nominal et les proportions. Pour plus de détails, consulter la section 11.

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, on devrait avoir au moins 10 observations. 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 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 coefficients de variation (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 du plan*, on a d'abord calculé les effets du plan pour une vaste gamme de caractéristiques, puis pour chaque tableau, choisi une valeur conservatrice parmi tous les effets du 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 effets du plan, les tailles d'échantillon et les comptes de population qui ont servi à produire les tableaux de la variabilité d'échantillonnage approximative de même que les tableaux, sont disponibles à l'Annexe E. 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 (p. ex., 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'écart-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 1,0 %.
- 4) Le coefficient de variation approximatif de l'estimation est donc 1,0 %. Par conséquent, selon
- 5) 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 0,6 %, est le coefficient de variation (exprimé en pourcentage) à employer.
- 5) Le coefficient de variation de l'estimation est donc 0,6 %. 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 1,5 % le CV de l'estimation 1 (exprimé en pourcentage) et à 1,5 % 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 :

$$\begin{aligned} \sigma_{\hat{d}} &= \sqrt{[(0,194)(0,015)]^2 + [(0,162)(0,015)]^2} \\ &= 0,004 \end{aligned}$$

- 3) Le coefficient de variation de \hat{d} est donné par $\sigma_{\hat{d}} / \hat{d} = 0,004/0,032 = 0,125$.
- 4) Le coefficient de variation approximatif de la différence entre les estimations est donc 12,5 % (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 sans réserve.

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 2,3 %.
- 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 1,0 %.
- 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,

$$\begin{aligned} \alpha_{\hat{R}} &= \sqrt{(0,023)^2 + (0,01)^2} \\ &= 0,025 \end{aligned}$$

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 2,5 % (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).

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

$$z = 2$$

$$\alpha \hat{x} = 0,006 \text{ est le coefficient de variation de cette estimation selon les tableaux.}$$

$$CI_{\hat{x}} = \{0,777 - (2) (0,777) (0,006), 0,777 + (2) (0,777) (0,006)\}$$

$$CI_{\hat{x}} = \{0,768, 0,786\}$$

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 $\sigma_{\hat{d}}$. Si $z = (\hat{X}_1 - \hat{X}_2) / \sigma_{\hat{d}}$ 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 0,004. Par conséquent,

$$z = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}} = \frac{0,194 - 0,162}{0,004} = \frac{0,032}{0,004} = 8$$

Puisque $z = 8$ 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 de 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, donc, jugée grossière, alors 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

L'Annexe E présente les tableaux indiquant les seuils de diffusion des totaux selon les estimations pour le Canada, les provinces, les régions sociosanitaires, les régions de CLSC ainsi que pour les différents groupes d'âges (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).

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, par télé-access et peuvent aussi faire l'objet de demandes de totalisations personnalisées.

12.1.1 Centres 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 Télé-access

Le service de télé-access 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 aux utilisateurs 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-escs@statcan.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.

12.1.3 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.ca.

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 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 FMGD sont accessibles gratuitement dans les établissements d'enseignement post-secondaires qui font partie de l'Initiative de démocratisation des données. Ils sont aussi disponibles gratuitement sur demande auprès du Service à la clientèle, pour tous les utilisateurs qui n'en font pas un usage commercial.

12.4 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.

Le module **Sécurité alimentaire** 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. La variable de pondération **WTS_M** doit être utilisée pour des analyses régionales.

12.5 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 = cycle 1.1, C = cycle 2.1 et E = 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 »)¹⁴.

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.

Tableau 12.1 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)

14. Du mot anglais *flag*.

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.

12.6 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 pour les cycles 1.1 à 3.1.

Voici la liste des lettres utilisées dans les fichiers de microdonnées de l'ESCC entre les cycles 1.1 et 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

12.7 Lignes directrices pour l'utilisation des variables d'un sous-échantillon

En 2007, un sous-échantillon de répondants a été choisi au hasard pour répondre aux questions de trois modules portant sur le rendement du système des soins de santé au Canada : Accès aux services de santé (ACC), Temps d'attente (WTM) et Satisfaction des patients (PAS)¹⁵.

12.7.1 Convention de dénomination des variables des modules de sous-échantillon

Les variables des modules de contenu de sous-échantillon sont désignées par la lettre « Z » à la position 4e du nom de la variable (par ex., ACCZ_01).

Chacun des modules du sous-échantillon a aussi été mis à la disposition des régions sociosanitaires à titre de contenu optionnel. Le module Accès aux services de santé a été choisi par les régions sociosanitaires de l'Ontario et du Nouveau-Brunswick, alors que le module Satisfaction des patients a été choisi par l'Ontario, la Saskatchewan, le Yukon et les Territoires du Nord-Ouest. Pour ces provinces, il existe deux ensembles de variables (par ex., ACC_11 et ACCZ_11). Les variables s'appliquent ou non à un répondant donné, selon qu'il fait partie ou non du sous-échantillon pertinent **et** qu'il vit ou non dans une région sociosanitaire où le module a été choisi comme contenu optionnel. Le tableau 1 illustre les possibilités.

Tableau 12.2 Valeurs potentiellement non applicables à une variable de sous-échantillon

Répondant du sous-échantillon	Répondant vivant dans une région sociosanitaire où le module Accès aux services de soins de santé (ACC) faisait partie du contenu optionnel	
	OUI	NON
OUI	ACC_11 et ACCZ_11 s'appliquent	ACC_11 ne s'applique pas et ACCZ_11 s'applique
NON	ACC_11 s'applique et ACCZ_11 ne s'applique pas	ACC_11 et ACCZ_11 ne s'appliquent pas

15. En 2000-2001, ce sous-échantillon était une enquête de suivi à l'ESCC connu sous le nom l'Enquête sur l'accès aux services de santé (EASS). L'enquête a été intégrée à l'ESCC en 2003.

12.7.2 Fichier distinct

Afin d'encourager l'utilisation appropriée des données de l'ESCC, le contenu du sous-échantillon est fourni dans un fichier distinct. Il comporte son propre poids d'échantillonnage et son ensemble de poids bootstrap correspondants, qui doivent être utilisés pour produire des estimations valables pour toutes les variables du fichier. Le poids d'échantillonnage du sous-échantillon 1 (dont les modules s'appliquent aux répondants âgés de 15 ans et plus) est calibré de façon à représenter les Canadiens âgés de 15 ans et plus. Le tableau 3 décrit les deux fichiers de données diffusés en 2007.

Tableau 12.3 Noms et contenu des fichiers de données de l'ESCC 2007

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 et optionnels	Enregistrements de tous les répondants
Fichier maître du sous-Échantillon 1	HSS1.txt	WTS_S1M	b5_s1.txt	Tous les modules communs, plus l'ensemble des variables « Z » pour le module Accès aux services de santé, Temps d'attente, Satisfaction des patients	Enregistrements de tous les répondants 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 le module Accès aux services de santé, Temps d'attente, Satisfaction des patients	Enregistrements de tous les répondants sélectionnés pour le sous-échantillon qui ont accepté de partager leurs données

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

Les modules du sous-échantillon de l'ESCC visent à fonder le calcul d'estimations au niveau national et provincial. 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.

Tel que mentionné précédemment, deux des trois modules du sous-échantillon ont également été choisis par l'ensemble des régions sociosanitaires d'au moins une province à titre de contenu optionnel. Dans ces cas, il nous est possible de calculer des estimations au niveau de la région sociosanitaire et de la province pour les variables du module en utilisant le **fichier maître principal** et le **poids d'échantillonnage maître**. Cela a pour avantage de fonder les calculs sur un échantillon important et donc de réduire l'erreur d'échantillonnage. L'estimation provinciale peut par la suite être comparée à l'estimation concernant d'autres provinces, calculé au moyen du fichier de sous-échantillon pertinent et du poids d'échantillonnage de sous-échantillon.

12.8 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 PASFOPT indique si le module Satisfaction des patients s'applique à un répondant donné.

12.9 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 fournis. 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

CONTENU DE L'ESCC POUR 2007 ET 2008

Contenu de base (toutes les régions)				
<ul style="list-style-type: none"> • Âge du répondant (ANC) • Consommation d'alcool (ALC) • Problèmes de santé chroniques (CCC) • Exposition à la fumée secondaire (ETS) • Vaccins contre la grippe (FLU) • État de santé général (GEN) • Utilisation des soins de santé (HCU) • Douleurs et malaise (HUP) 		<ul style="list-style-type: none"> • Taille et poids – autodéclarés (HWT) • Expériences maternelles – Allaitement (MEX) • Consommation de fruits et de légumes (FVC) • Activités physiques (PAC) • Limitation des activités (RAC) • Usage du tabac (SMK) 		<p>Administration et renseignements sociodémographiques</p> <ul style="list-style-type: none"> • Renseignements administratifs (ADM) • Renseignement sur le logement (DWL) • Éducation (EDU) • Revenu (INC) • Population active (LBS) <p>Renseignements sociodémographiques</p>
Contenu thématique (toutes les régions)				
<i>Thème pour 2007-2008: Mode de vie sain</i>		<i>Thème pour 2007: Accès aux services de soins de santé au Canada i</i>		<i>Thème pour 2008: Prévention des maladies chroniques</i>
<ul style="list-style-type: none"> • Changements faits pour améliorer la santé (CIH) • Sécurité alimentaire (FSC) • Santé bucco-dentaire 1 (OH1) 	<ul style="list-style-type: none"> • Activités physiques – Installations au travail (PAF) • Activités sédentaires (SAC) 	<ul style="list-style-type: none"> • Accès aux services de soins de santé (ACCZ) • Satisfaction des patients à l'égard des services de soins de santé (PASZ) • Temps d'attente (WTMZ) 	<ul style="list-style-type: none"> • Analyse de sang (BLT) • Test pap (PAP) • Mammographie (MAM) 	<ul style="list-style-type: none"> • Spirométrie (SPI) • Examen général (PCU) • Dépistage du cancer du côlon et du rectum (CCS)

Contenu optionnel (certaines régions)			
<ul style="list-style-type: none"> • Accès aux services de soins de santé (ACC)ⁱⁱ • Consommation d'alcool – Dépendance (ALD) • Consommation d'alcool – Anciens buveurs (ALN) • Consommation d'alcool au cours de la dernière semaine (ALW) • Tension artérielle – Vérification (BPC) • Examen des seins (BRX) • Auto-examen des seins (BSX) • Stress - Facteurs de stress durant l'enfance et la vie adulte (CST) • 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) 	<ul style="list-style-type: none"> • Examens de la vue (EYX) • Choix alimentaires (FDC) • Indice de l'état de santé (HUI) • État de santé (SFR) • Services de soins de santé à domicile (HMC) • Mesures de sécurité à la maison (HMS) • Usage de drogues illicites (DRG) • Blessures (INJ) • 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) • Usage du tabac - Dépendance à la nicotine (NDE) 	<ul style="list-style-type: none"> • Santé bucco-dentaire 2 (OH2) • Test pap (PAP) • Satisfaction des patients à l'égard des services de soins de santé (PAS)ⁱⁱ • Satisfaction des patients à l'égard des soins de santé communautaires (PSC) • Jeu excessif (CPG) • Dépistage du cancer de la prostate (CCS) • Bien-être psychologique (PWB) • Stress - Événements récents (RLE) • Satisfaction à l'égard de la vie (SWL) • Estime de soi (SFE) • Comportements sexuels (SXB) • Sommeil (SLP) • Usage du tabac - Méthodes pour cesser de fumer (SCA) • Usage du tabac – Consultation d'un médecin (SPC) 	<ul style="list-style-type: none"> • Usage du tabac – Les étapes du changement (SCH) • Soutien social – Disponibilité (SSA) • Soutien social – Utilisation (SSU) • Valeurs spirituelles (SPR) • 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) • Utilisation de l'équipement protecteur (UPE) • Organismes à but non lucratif – Participation (ORG) • Stress au travail (WST) • Usage du tabac chez les jeunes (YSM)

i Demandé à un sous-échantillon de répondants. Le thème de 2007 (Accès aux services de soins de santé au Canada) n'est pas demandé aux répondants des territoires.

ii Ces modules de contenu thématique de 2007 ont aussi été sélectionnés comme contenu optionnel par certaines régions

ANNEXE B

**SÉLECTION DU CONTENU OPTIONNEL SELON
LA PROVINCE ET LE TERRITOIRE, 2007**

Modules optionnels	T.-N.-L.	Î.-P.-É.	N.-É.	N.-B.	Qc	Ont.	Man.	Sask.	Alb.	C.-B.	Yn	T.N.-O.	Nt
Accès aux services de soins de santé				*		*							
Consommation d'alcool - Dépendance								*		*			
Consommation d'alcool – Anciens buveurs	*									*			
Consommation d'alcool au cours de la dernière semaine	*		*			*				*			*
Tension artérielle - Vérification								*	*			*	
Examen des seins				*				*				*	
Auto-examen des seins	*			*				*					
Stress - Facteurs de stress durant l'enfance et la vie adulte												*	
Dépistage du cancer du côlon et du rectum	*	*				*		*				*	*
Consultations au sujet de la santé mentale		*	*	*		*			*			*	*
Visites chez le dentiste	*								*			*	
Dépression			*	*	*				*		*		*
Soins pour le diabète	*	*	*	*		*				*	*	*	
Usage de suppléments alimentaires - Vitamines et minéraux											*	*	
Détresse			*		*								
Conduite et sécurité							*		*			*	
Examens de la vue						*							
Choix alimentaires		*					*		*	*		*	
Satisfaction à l'égard du système de soins de santé						*					*	*	
Indice de l'état de santé					*								
Services de soins de santé à domicile						*							
Mesures de sécurité à la maison												*	
Usage de drogues illicites			*							*			
Blessures			*							*			
Couverture d'assurance				*									

Modules optionnels	T.-N.-L.	Î.-P.-É.	N.-É.	N.-B.	Qc	Ont.	Man.	Sask.	Alb.	C.-B.	Yn	T.N.-O.	Nt
Mammographie	*		*	*		*		*				*	
Contrôle de soi							*						
Expériences maternelles – Consommation d'alcool au cours de la grossesse						*		*		*		*	
Expériences maternelles – Usage du tabac au cours de la grossesse						*		*		*		*	*
Usage du tabac - Dépendance à la nicotine									*			*	*
Santé bucco-dentaire 2	*												
Test pap				*		*		*				*	
Satisfaction des patients à l'égard des services de soins de santé						*		*			*	*	
Satisfaction des patients à l'égard des soins de santé communautaires				*				*			*	*	
Jeu excessif					*	*		*					*
Dépistage du cancer de la prostate	*	*						*				*	
Bien-être psychologique				*									
Stress - Événements récents											*		
Satisfaction à l'égard de la vie					*				*				*
Estime de soi													*
Comportements sexuels			*	*		*		*				*	*
État de santé (SF-36)	*	*					*						
Sommeil			*					*			*		
Usage du tabac – Consultation d'un médecin									*				
Usage du tabac – Les étapes du changement						*							
Usage du tabac - Méthodes pour cesser de fumer						*							
Soutien social - Disponibilité			*		*					*	*		*
Soutien social - Utilisation										*	*		*

Modules optionnels	T.-N.-L.	Î.-P.-É.	N.-É.	N.-B.	Qc	Ont.	Man.	Sask.	Alb.	C.-B.	Yn	T.N.-O.	Nt
Valeurs spirituelles								*					*
Stress - Faire face au stress											*		*
Stress - Sources											*		*
Pensées suicidaires et tentatives de suicide						*			*	*		*	
Protection contre le soleil		*											
Usage du tabac - Autres produits du tabac						*		*	*				
Utilisation de l'équipement protecteur			*					*	*				
Organismes à but non lucratif - Participation			*									*	
Stress au travail									*				
Usage du tabac chez les jeunes										*			

ANNEXE C

GÉOGRAPHIE DE L'ESCC

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
		Réseau d'intégration des services de santé de Hamilton Niagara
3504		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		Norman 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	Territoire du Yukon	
6001-H		Territoire du 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

**RÉPARTITION DE L'ÉCHANTILLON PAR RÉGION
SOCIOSANITAIRE ET PAR BASE DE SONDAGE**

Répartition de l'échantillon par région socio-sanitaire et par base de sondage de l'ESCC 2007

Prov./Terr.	Région socio-sanitaire	Base aréolaire		Bases téléphoniques		/Combiné	
		# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute
Canada	Total	33375	47120	32903	53978	66278	101098
T.-N.-L.	Total	1003	1397	1003	1353	2006	2750
	1011	405	556	405	548	810	1104
	1012	235	340	235	324	470	664
	1013	213	271	213	282	426	553
	1014	150	230	150	199	300	429
I.-P.-É.	Total	741	1110	739	1134	1480	2244
	1101	88	150	88	138	176	288
	1102	410	618	418	648	828	1266
	1103	243	342	233	348	476	690
N.-É.	Total	1261	1864	1261	1748	2522	3612
	1201	198	300	198	282	396	582
	1202	160	219	160	212	320	431
	1203	180	239	180	261	360	500
	1204	175	300	175	257	350	557
	1205	210	297	210	286	420	583
	1206	338	509	338	450	676	959
N.-B.	Total	1289	1811	1289	2115	2578	3926
	1301	250	358	250	413	500	771
	1302	243	361	243	403	486	764
	1303	235	340	235	391	470	731
	1304	135	183	135	232	270	415
	1305	125	172	125	209	250	381
	1306	173	232	173	251	346	483
	1307	128	165	128	216	256	381
Qc	Total	5872	8161	6272	10753	12144	18914
	2401	300	395	300	449	600	844
	2402	314	402	314	453	628	855
	2403	464	608	464	711	928	1319
	2404	401	523	401	601	802	1124
	2405	309	470	309	471	618	941
	2406	776	1110	776	1357	1552	2467
	2407	323	485	323	504	646	989
	2408	300	436	300	429	600	865
	2409	300	423	300	457	600	880
	2410	0	0	400	1584	400	1584
	2411	300	418	300	452	600	870
	2412	361	480	361	556	722	1036
	2413	335	466	335	540	670	1006
	2414	359	472	359	570	718	1042
	2415	380	604	380	633	760	1237
	2416	650	869	650	986	1300	1855
Ont.	Total	11114	15657	11114	18290	22228	33947

Prov./Terr.	Région socio-sanitaire	Base aréolaire		Bases téléphoniques		/Combiné	
		# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute
	3526	213	342	213	338	426	680
	3527	203	284	203	315	406	599
	3530	408	582	408	629	816	1211
	3531	170	222	170	275	340	497
	3533	237	340	237	401	474	741
	3534	193	283	193	331	386	614
	3535	238	391	238	434	476	825
	3536	353	445	353	544	706	989
	3537	413	592	413	663	826	1255
	3538	235	336	235	389	470	725
	3539	148	186	148	231	296	417
	3540	200	243	200	311	400	554
	3541	253	390	253	410	506	800
	3542	218	304	218	360	436	664
	3543	238	337	238	380	476	717
	3544	375	582	375	597	750	1179
	3546	383	497	383	623	766	1120
	3547	200	281	200	325	400	606
	3549	164	258	164	336	328	594
	3551	513	746	513	774	1026	1520
	3552	188	258	188	256	376	514
	3553	657	896	657	1130	1314	2026
	3554	163	221	163	241	326	462
	3555	213	327	213	380	426	707
	3556	188	264	188	297	376	561
	3557	188	271	188	326	376	597
	3558	260	347	260	387	520	734
	3560	571	844	571	947	1142	1791
	3561	270	359	270	437	540	796
	3562	358	486	358	589	716	1075
	3563	125	186	125	196	250	382
	3565	383	515	383	560	766	1075
	3566	282	367	282	431	564	798
	3568	358	481	358	566	716	1047
	3570	471	604	471	814	942	1418
	3595	1084	1590	1084	2067	2168	3657
Man.	Total	1877	2552	1877	2829	3754	5381
	4610	528	705	528	747	1056	1452
	4615	140	185	140	217	280	402
	4620	125	180	125	240	250	420
	4625	150	199	150	221	300	420
	4630	173	280	173	263	346	543
	4640	200	251	200	272	400	523
	4645	178	240	178	252	356	492
	4660	133	189	133	193	266	382

Prov./Terr.	Région socio-sanitaire	Base aréolaire		Bases téléphoniques		/Combiné	
		# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute
	4670	125	169	125	224	250	393
	4685	125	154	125	200	250	354
Sask.	Total	1806	2510	2056	3890	3862	6400
	4701	150	193	150	220	300	413
	4702	150	213	150	228	300	441
	4703	133	187	133	198	266	385
	4704	310	427	310	436	620	863
	4705	155	219	155	229	310	448
	4706	330	443	330	474	660	917
	4707	135	227	135	188	270	415
	4708	130	191	130	180	260	371
	4709	163	217	163	259	326	476
	4710	150	193	150	227	300	420
	4714	0	0	250	1251	250	1251
Alb.	Total	3052	4244	3052	4642	6104	8886
	4820	255	338	255	381	510	719
	4821	208	255	208	300	416	555
	4822	698	1019	698	1036	1396	2055
	4823	350	503	350	529	700	1032
	4824	223	297	223	333	446	630
	4825	655	904	655	963	1310	1867
	4826	270	415	270	445	540	860
	4827	233	312	233	376	466	688
	4828	160	201	160	279	320	480
C.-B.	Total	4025	5707	4025	6625	8050	12332
	5911	152	205	152	245	304	450
	5912	155	211	155	244	310	455
	5913	295	356	295	479	590	835
	5914	250	380	250	408	500	788
	5921	260	346	260	412	520	758
	5922	380	471	380	640	760	1111
	5923	400	568	400	657	800	1225
	5931	213	293	213	367	426	660
	5932	400	646	400	750	800	1396
	5933	273	411	273	447	546	858
	5941	338	468	338	534	676	1002
	5942	263	365	263	399	526	764
	5943	132	182	132	200	264	382
	5951	163	275	163	293	326	568
	5952	213	330	213	323	426	653
	5903	138	200	138	227	276	427
Terr.	60	475	733	125	334	600	1067
	61	510	815	90	265	600	1080
	62	350	559	0	0	350	559

**Répartition de l'échantillon par réseau local d'intégration d'intégration des services de santé
et par base de sondage de l'ESCC 2007 en Ontario**

		Base aréolaire		Bases téléphoniques		Combiné	
Prov./Terr.	RLISS	# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute	# répondants espérés	taille d'échantillon brute
Ont.	Total	11114	15657	11114	18290	22228	33947
	3501	776	1028	776	1237	1552	2265
	3502	1284	1815	1284	2008	2568	3823
	3503	615	818	615	916	1230	1734
	3504	1310	1800	1310	2110	2620	3910
	3505	531	729	531	921	1062	1650
	3506	568	756	568	945	1136	1701
	3507	542	796	542	1034	1084	1830
	3508	702	948	702	1247	1404	2195
	3509	1038	1560	1038	1786	2076	3346
	3510	672	990	672	1096	1344	2086
	3511	1029	1460	1029	1595	2058	3055
	3512	529	781	529	877	1058	1658
	3513	996	1432	996	1593	1992	3025
	3514	522	744	522	925	1044	1669

**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) 2007**

		Base aréolaire		Bases téléphoniques		Combiné	
Prov./Terr.	# répondants espérés ¹	taille d'échantillon brute	# répondants espérés ¹	taille d'échantillon brute	# répondants espérés ¹	taille d'échantillon brute	
Canada	23,690	35,789	23,690	40,844	47,380	76,633	
T.-N.-L.	937	1,397	937	1,353	1,874	2,750	
Î.-P.-É.	693	1,110	693	1,134	1,386	2,244	
N.-É.	1,178	1,864	1,178	1,748	2,356	3,612	
N.-B.	1,204	1,811	1,204	2,115	2,408	3,926	
Qc	2,300	3,449	2,300	4,003	4,600	7,452	
Ont.	10,378	15,657	10,378	18,290	20,756	33,947	
Man.	1,600	2,363	1,600	2,598	3,200	4,961	
Sask.	1,600	2,325	1,600	3,071	3,200	5,396	
Alb.	1,800	2,752	1,800	2,944	3,600	5,696	
C.-B.	2,000	3,061	2,000	3,588	4,000	6,649	

ANNEXE E

TAUX DE RÉPONSE PAR RÉGION SOCIO-SANITAIRE ET PAR BASE DE SONDAGE

Taux de réponse par région socio-sanitaire et par base de sondage de l'ESCC 2007

		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
Canada	Total	40663	35748	87.9	35748	33473	93.6	82.3	44310	36174	81.6	36174	32473	89.8	73.3	77.6
T.-N.-L.	Total	1192	1111	93.2	1111	1036	93.2	86.9	1247	1065	85.4	1065	958	90	76.8	81.8
	1011	491	453	92.3	453	426	94	86.8	488	415	85	415	370	89.2	75.8	81.3
	1012	269	257	95.5	257	240	93.4	89.2	323	277	85.8	277	252	91	78	83.1
	1013	227	217	95.6	217	210	96.8	92.5	252	215	85.3	215	194	90.2	77	84.3
	1014	205	184	89.8	184	160	87	78	184	158	85.9	158	142	89.9	77.2	77.6
Î.-P.-É.	Total	889	790	88.9	790	744	94.2	83.7	888	733	82.5	733	659	89.9	74.2	79
	1101	82	70	85.4	70	65	92.9	79.3	61	50	82	50	44	88	72.1	76.2
	1102	553	481	87	481	451	93.8	81.6	567	466	82.2	466	429	92.1	75.7	78.6
	1103	254	239	94.1	239	228	95.4	89.8	260	217	83.5	217	186	85.7	71.5	80.5
N.-É.	Total	1543	1368	88.7	1368	1270	92.8	82.3	1665	1415	85	1415	1304	92.2	78.3	80.2
	1201	219	208	95	208	199	95.7	90.9	261	225	86.2	225	210	93.3	80.5	85.2
	1202	191	174	91.1	174	159	91.4	83.2	188	160	85.1	160	148	92.5	78.7	81
	1203	215	196	91.2	196	190	96.9	88.4	232	194	83.6	194	184	94.8	79.3	83.7
	1204	214	182	85	182	163	89.6	76.2	227	189	83.3	189	171	90.5	75.3	75.7
	1205	253	235	92.9	235	222	94.5	87.7	274	233	85	233	213	91.4	77.7	82.5
	1206	451	373	82.7	373	337	90.3	74.7	483	414	85.7	414	378	91.3	78.3	76.6
N.-B.	Total	1518	1358	89.5	1358	1277	94	84.1	1792	1540	85.9	1540	1427	92.7	79.6	81.7
	1301	305	266	87.2	266	242	91	79.3	350	294	84	294	275	93.5	78.6	78.9
	1302	278	244	87.8	244	232	95.1	83.5	338	286	84.6	286	265	92.7	78.4	80.7
	1303	271	239	88.2	239	226	94.6	83.4	318	271	85.2	271	250	92.3	78.6	80.8
	1304	169	141	83.4	141	136	96.5	80.5	183	162	88.5	162	153	94.4	83.6	82.1
	1305	150	138	92	138	132	95.7	88	186	166	89.2	166	148	89.2	79.6	83.3
	1306	199	192	96.5	192	178	92.7	89.4	231	204	88.3	204	192	94.1	83.1	86

		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	146	138	94.5	138	131	94.9	89.7	186	157	84.4	157	144	91.7	77.4	82.8
Qc	Total	7202	6338	88	6338	6017	94.9	83.5	8506	6960	81.8	6960	6233	89.6	73.3	78
	2401	344	320	93	320	314	98.1	91.3	364	321	88.2	321	294	91.6	80.8	85.9
	2402	353	321	90.9	321	300	93.5	85	382	323	84.6	323	291	90.1	76.2	80.4
	2403	606	531	87.6	531	493	92.8	81.4	651	534	82	534	483	90.4	74.2	77.6
	2404	479	429	89.6	429	411	95.8	85.8	555	465	83.8	465	422	90.8	76	80.6
	2405	404	336	83.2	336	320	95.2	79.2	370	324	87.6	324	300	92.6	81.1	80.1
	2406	969	820	84.6	820	772	94.1	79.7	1248	904	72.4	904	777	86	62.3	69.9
	2407	416	341	82	341	323	94.7	77.6	431	352	81.7	352	320	90.9	74.2	75.9
	2408	382	359	94	359	347	96.7	90.8	346	311	89.9	311	291	93.6	84.1	87.6
	2409	373	330	88.5	330	315	95.5	84.5	383	327	85.4	327	292	89.3	76.2	80.3
	2410	541	433	80	433	386	89.1	71.3	71.3
	2411	342	322	94.2	322	311	96.6	90.9	374	327	87.4	327	283	86.5	75.7	83
	2412	404	376	93.1	376	364	96.8	90.1	457	390	85.3	390	360	92.3	78.8	84.1
	2413	429	368	85.8	368	333	90.5	77.6	539	424	78.7	424	362	85.4	67.2	71.8
	2414	444	383	86.3	383	369	96.3	83.1	476	387	81.3	387	352	91	73.9	78.4
	2415	445	401	90.1	401	375	93.5	84.3	509	414	81.3	414	374	90.3	73.5	78.5
	2416	812	701	86.3	701	670	95.6	82.5	880	724	82.3	724	646	89.2	73.4	77.8
Ont.	Total	13841	12031	86.9	12031	11183	93	80.8	15419	12251	79.5	12251	10784	88	69.9	75.1
	3526	298	249	83.6	249	229	92	76.8	278	238	85.6	238	217	91.2	78.1	77.4
	3527	256	231	90.2	231	205	88.7	80.1	269	219	81.4	219	201	91.8	74.7	77.3
	3530	521	454	87.1	454	418	92.1	80.2	559	448	80.1	448	393	87.7	70.3	75.1
	3531	193	169	87.6	169	157	92.9	81.3	217	176	81.1	176	161	91.5	74.2	77.6
	3533	269	242	90	242	233	96.3	86.6	308	248	80.5	248	227	91.5	73.7	79.7
	3534	238	208	87.4	208	188	90.4	79	275	218	79.3	218	194	89	70.5	74.5
	3535	344	299	86.9	299	293	98	85.2	326	269	82.5	269	241	89.6	73.9	79.7

		Base aéroilaire							Bases téléphoniques							
3536		425	362	85.2	362	337	93.1	79.3	476	388	81.5	388	340	87.6	71.4	75.1
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	528	441	83.5	441	395	89.6	74.8	557	437	78.5	437	388	88.8	69.7	72.2
	3538	298	266	89.3	266	249	93.6	83.6	304	249	81.9	249	213	85.5	70.1	76.7
	3539	170	163	95.9	163	159	97.5	93.5	210	167	79.5	167	148	88.6	70.5	80.8
	3540	219	211	96.3	211	200	94.8	91.3	286	224	78.3	224	195	87.1	68.2	78.2
	3541	344	288	83.7	288	270	93.8	78.5	316	259	82	259	231	89.2	73.1	75.9
	3542	229	214	93.4	214	198	92.5	86.5	274	220	80.3	220	198	90	72.3	78.7
	3543	276	250	90.6	250	236	94.4	85.5	305	262	85.9	262	238	90.8	78	81.6
	3544	516	403	78.1	403	386	95.8	74.8	530	426	80.4	426	369	86.6	69.6	72.2
	3546	433	395	91.2	395	377	95.4	87.1	513	403	78.6	403	365	90.6	71.2	78.4
	3547	239	207	86.6	207	182	87.9	76.2	269	224	83.3	224	204	91.1	75.8	76
	3549	194	178	91.8	178	155	87.1	79.9	212	169	79.7	169	140	82.8	66	72.7
	3551	674	559	82.9	559	512	91.6	76	670	548	81.8	548	486	88.7	72.5	74.3
	3552	240	222	92.5	222	205	92.3	85.4	226	195	86.3	195	173	88.7	76.5	81.1
	3553	854	745	87.2	745	675	90.6	79	971	736	75.8	736	616	83.7	63.4	70.7
	3554	206	186	90.3	186	169	90.9	82	215	185	86	185	164	88.6	76.3	79.1
	3555	249	212	85.1	212	201	94.8	80.7	272	227	83.5	227	208	91.6	76.5	78.5
	3556	237	205	86.5	205	183	89.3	77.2	237	206	86.9	206	194	94.2	81.9	79.5
	3557	216	199	92.1	199	182	91.5	84.3	242	197	81.4	197	185	93.9	76.4	80.1
	3558	321	277	86.3	277	264	95.3	82.2	348	292	83.9	292	259	88.7	74.4	78.2
	3560	746	642	86.1	642	597	93	80	828	684	82.6	684	600	87.7	72.5	76
	3561	311	273	87.8	273	259	94.9	83.3	349	289	82.8	289	264	91.3	75.6	79.2
	3562	410	360	87.8	360	340	94.4	82.9	491	407	82.9	407	363	89.2	73.9	78
	3563	125	118	94.4	118	116	98.3	92.8	166	132	79.5	132	116	87.9	69.9	79.7
	3565	472	392	83.1	392	359	91.6	76.1	489	407	83.2	407	371	91.2	75.9	76
	3566	349	317	90.8	317	301	95	86.2	390	321	82.3	321	293	91.3	75.1	80.4

		Base aréolaire							Bases téléphoniques							
	3568	416	366	88	366	347	94.8	83.4	507	412	81.3	412	346	84	68.2	75.1
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
	3570	566	508	89.8	508	477	93.9	84.3	732	533	72.8	533	449	84.2	61.3	71.3
	3595	1459	1220	83.6	1220	1129	92.5	77.4	1802	1236	68.6	1236	1034	83.7	57.4	66.3
Man.	Total	2202	1996	90.6	1996	1895	94.9	86.1	2261	1968	87	1968	1831	93	81	83.5
	4610	652	568	87.1	568	541	95.2	83	655	550	84	550	512	93.1	78.2	80.6
	4615	165	148	89.7	148	137	92.6	83	172	142	82.6	142	130	91.5	75.6	79.2
	4620	158	143	90.5	143	135	94.4	85.4	175	157	89.7	157	151	96.2	86.3	85.9
	4625	167	158	94.6	158	147	93	88	187	167	89.3	167	149	89.2	79.7	83.6
	4630	203	186	91.6	186	180	96.8	88.7	205	184	89.8	184	167	90.8	81.5	85
	4640	228	212	93	212	200	94.3	87.7	223	201	90.1	201	191	95	85.7	86.7
	4645	205	190	92.7	190	180	94.7	87.8	203	184	90.6	184	174	94.6	85.7	86.8
	4660	145	133	91.7	133	132	99.2	91	161	139	86.3	139	131	94.2	81.4	85.9
	4670	134	128	95.5	128	117	91.4	87.3	149	127	85.2	127	121	95.3	81.2	84.1
	4685	145	130	89.7	130	126	96.9	86.9	131	117	89.3	117	105	89.7	80.2	83.7
Sask.	Total	2068	1910	92.4	1910	1810	94.8	87.5	2579	2191	85	2191	2008	91.6	77.9	82.2
	4701	174	167	96	167	160	95.8	92	165	148	89.7	148	137	92.6	83	87.6
	4702	177	166	93.8	166	156	94	88.1	189	154	81.5	154	140	90.9	74.1	80.9
	4703	159	143	89.9	143	133	93	83.6	171	131	76.6	131	121	92.4	70.8	77
	4704	389	352	90.5	352	333	94.6	85.6	435	366	84.1	366	332	90.7	76.3	80.7
	4705	154	148	96.1	148	143	96.6	92.9	168	141	83.9	141	136	96.5	81	86.6
	4706	387	351	90.7	351	330	94	85.3	448	375	83.7	375	340	90.7	75.9	80.2
	4707	154	145	94.2	145	139	95.9	90.3	162	135	83.3	135	120	88.9	74.1	82
	4708	148	141	95.3	141	134	95	90.5	130	111	85.4	111	101	91	77.7	84.5
	4709	151	131	86.8	131	123	93.9	81.5	248	215	86.7	215	197	91.6	79.4	80.2
	4710	175	166	94.9	166	159	95.8	90.9	166	145	87.3	145	138	95.2	83.1	87.1
	4714	297	270	90.9	270	246	91.1	82.8	82.8

		Base aréolaire							Bases téléphoniques							
Alb.	Total	3669	3212	87.5	3212	2961	92.2	80.7	4097	3363	82.1	3363	3032	90.2	74	77.2
	4821	297	270	90.9	270	244	90.4	82.2	319	257	80.6	257	242	94.2	75.9	78.9
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	213	198	93	198	190	96	89.2	268	232	86.6	232	210	90.5	78.4	83.2
	4823	904	764	84.5	764	704	92.1	77.9	917	760	82.9	760	688	90.5	75	76.4
	4824	454	395	87	395	363	91.9	80	466	396	85	396	356	89.9	76.4	78.2
	4825	260	246	94.6	246	231	93.9	88.8	301	256	85	256	238	93	79.1	83.6
	4826	796	692	86.9	692	639	92.3	80.3	953	750	78.7	750	661	88.1	69.4	74.3
	4827	309	281	90.9	281	264	94	85.4	337	277	82.2	277	246	88.8	73	78.9
	4828	274	239	87.2	239	213	89.1	77.7	316	259	82	259	233	90	73.7	75.6
	4829	162	127	78.4	127	113	89	69.8	220	176	80	176	158	89.8	71.8	70.9
C.-B.	Total	4857	4130	85	4130	3863	93.5	79.5	5591	4451	79.6	4451	4026	90.5	72	75.5
	5911	167	161	96.4	161	155	96.3	92.8	221	173	78.3	173	160	92.5	72.4	81.2
	5912	160	137	85.6	137	130	94.9	81.3	187	146	78.1	146	137	93.8	73.3	76.9
	5913	356	308	86.5	308	296	96.1	83.1	447	359	80.3	359	326	90.8	72.9	77.5
	5914	289	266	92	266	255	95.9	88.2	307	257	83.7	257	229	89.1	74.6	81.2
	5921	299	274	91.6	274	261	95.3	87.3	339	291	85.8	291	270	92.8	79.6	83.2
	5922	433	355	82	355	342	96.3	79	538	429	79.7	429	394	91.8	73.2	75.8
	5923	497	422	84.9	422	382	90.5	76.9	578	456	78.9	456	395	86.6	68.3	72.3
	5931	256	226	88.3	226	218	96.5	85.2	312	243	77.9	243	220	90.5	70.5	77.1
	5932	556	426	76.6	426	407	95.5	73.2	627	462	73.7	462	415	89.8	66.2	69.5
	5933	292	253	86.6	253	224	88.5	76.7	373	285	76.4	285	259	90.9	69.4	72.6
	5941	426	360	84.5	360	331	91.9	77.7	426	348	81.7	348	324	93.1	76.1	76.9
	5942	305	263	86.2	263	246	93.5	80.7	341	284	83.3	284	264	93	77.4	78.9
	5943	175	163	93.1	163	155	95.1	88.6	179	152	84.9	152	137	90.1	76.5	82.5
	5951	214	163	76.2	163	149	91.4	69.6	226	170	75.2	170	143	84.1	63.3	66.4
	5952	317	277	87.4	277	246	88.8	77.6	306	243	79.4	243	215	88.5	70.3	74

		Base aréolaire							Bases téléphoniques							
	5953	115	76	66.1	76	66	86.8	57.4	184	153	83.2	153	138	90.2	75	68.2
Yn	6001	615	531	86.3	531	500	94.2	81.3	160	146	91.3	146	129	88.4	80.6	81.2
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	632	583	92.2	583	543	93.1	85.9	105	91	86.7	91	82	90.1	78.1	84.8
Nt	6201	435	390	89.7	390	374	95.9	86	86

Taux de réponse par réseau local d'intégration des services de santé (RLISS) et par base de sondage de l'ESCC 2007 en Ontario

		Area frame / Base aréolaire							Phone frames / 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	13841	12031	86.9	12031	11183	93	80.8	15419	12251	79.5	12251	10784	88	69.9	75.1
	3501	864	791	91.6	791	745	94.2	86.2	1067	856	80.2	856	739	86.3	69.3	76.9
	3502	1592	1387	87.1	1387	1310	94.4	82.3	1699	1390	81.8	1390	1238	89.1	72.9	77.4
	3503	767	662	86.3	662	614	92.7	80.1	812	668	82.3	668	612	91.6	75.4	77.6
	3504	1599	1394	87.2	1394	1279	91.8	80	1774	1406	79.3	1406	1263	89.8	71.2	75.4
	3505	682	583	85.5	583	536	91.9	78.6	794	611	77	611	501	82	63.1	70.3
	3506	723	639	88.4	639	582	91.1	80.5	824	646	78.4	646	550	85.1	66.7	73.2
	3507	705	562	79.7	562	514	91.5	72.9	865	587	67.9	587	509	86.7	58.8	65.2
	3508	895	795	88.8	795	747	94	83.5	1157	811	70.1	811	682	84.1	58.9	69.6
	3509	1390	1209	87	1209	1141	94.4	82.1	1475	1167	79.1	1167	1021	87.5	69.2	75.5
	3510	829	721	87	721	678	94	81.8	830	693	83.5	693	609	87.9	73.4	77.6
	3511	1300	1118	86	1118	1035	92.6	79.6	1355	1114	82.2	1114	1003	90	74	76.8

	3512	681	580	85.2	580	538	92.8	79	765	637	83.3	637	559	87.8	73.1	75.9
	3513	1210	1052	86.9	1052	969	92.1	80.1	1299	1089	83.8	1089	995	91.4	76.6	78.3
	3514	604	538	89.1	538	495	92	82	703	576	81.9	576	503	87.3	71.6	76.4

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) 2007

Prov. Terr.	Base aréolaire							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	31046	27295	87.9	25968	24171	93.1	81.8	34553	28149	81.5	26136	22486	86	70.1	75.7
T.-N.-L.	1192	1111	93.2	1058	988	93.4	87	1247	1065	85.4	981	862	87.9	75	80.9
Î.-P.-É.	889	790	88.9	761	716	94.1	83.6	888	733	82.5	669	592	88.5	73	78.4
N.-É.	1543	1368	88.7	1317	1220	92.6	82.1	1665	1415	85	1302	1189	91.3	77.6	79.8
N.-B.	1518	1358	89.5	1288	1209	93.9	84	1792	1540	85.9	1409	1284	91.1	78.3	80.9
Qc	3076	2699	87.7	2569	2427	94.5	82.9	3519	2832	80.5	2656	2258	85	68.4	75.2
Ont.	13841	12031	86.9	11444	10592	92.6	80.5	15419	12251	79.5	11470	9634	84	66.7	73.3
Man.	2039	1842	90.3	1741	1641	94.3	85.1	2085	1808	86.7	1675	1479	88.3	76.6	80.8
Sask.	1923	1773	92.2	1673	1574	94.1	86.7	2253	1911	84.8	1774	1567	88.3	74.9	80.4
Alb.	2383	2063	86.6	1958	1783	91.1	78.8	2637	2156	81.8	1997	1733	86.8	71	74.7
C.-B.	2642	2260	85.5	2159	2021	93.6	80.1	3048	2438	80	2203	1888	85.7	68.5	74

