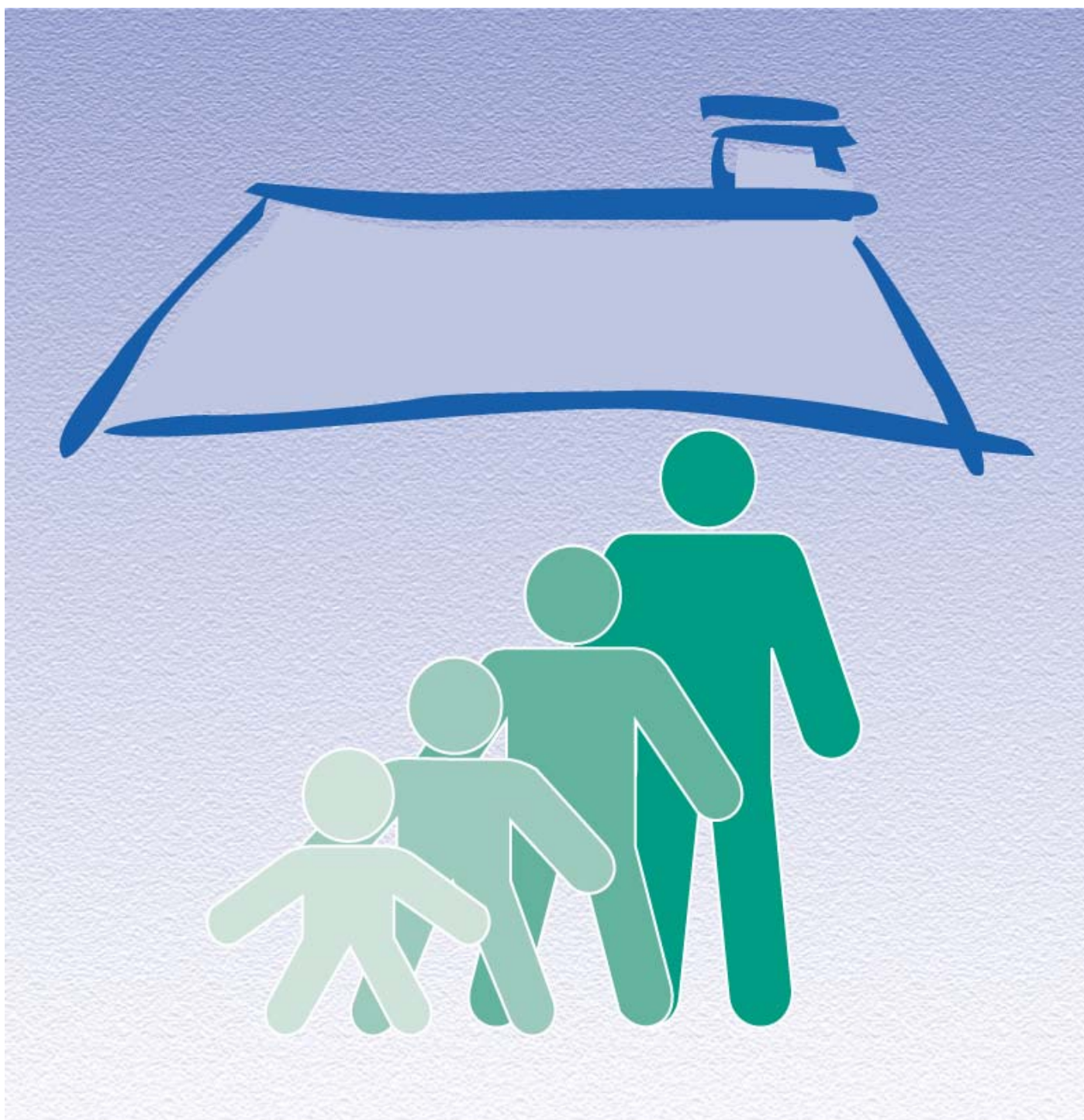


Enquête sur la santé dans les collectivités canadiennes (ESCC)

Documentation sur les poids ménages - 2007

Juin 2008



Statistique
Canada

Statistics
Canada

Canada

TABLE DES MATIÈRES

1.	INTRODUCTION	1
2.	VUE D'ENSEMBLE DE LA PONDÉRATION	1
3.	PONDÉRATION DE L'ÉCHANTILLON.....	2
3.1	NON-RÉPONSE PERSONNE (I2)	3
3.2	WINSORIZATION (I3)	3
3.3	CALAGE AUX MARGES (I4)	4
3.4	PARTICULARITÉS DE LA PONDÉRATION POUR LES TROIS TERRITOIRES	4

1. Introduction

Ce document décrit la stratégie de pondération pour la production du poids des ménages de 2007 de l'Enquête sur la Santé dans les Collectivités Canadiennes (ESCC). En utilisant le poids des ménages, les utilisateurs doivent être conscients que l'enquête a été conçue afin de représenter les individus et non pas les ménages. Le plan de sondage a été établi pour s'assurer que l'échantillon était représentatif de différents groupes démographiques au niveau des personnes. Il est possible que la qualité de la représentativité de l'échantillon par rapport à certains types de ménages soit moindre. De plus, puisque le calage aux marges pour le poids des ménages est effectué au niveau des provinces, il est possible de produire des estimations fiables aux échelles nationale et provinciale seulement. D'autre part, il peut être tentant d'utiliser les réponses des individus à certaines questions comme si elles représentaient le ménage alors que ce n'est pas le cas. Les utilisateurs doivent remarquer que le poids des ménages devrait être utilisé seulement pour les variables où on peut clairement affirmer que les réponses de l'individu représentent le ménage et que la réponse resterait la même si n'importe quel autre individu du ménage avait répondu. Ceci est appuyé par le fait que, tout au long du document, on fait référence aux personnes répondantes. Les utilisateurs doivent donc se souvenir qu'il est sous-entendu que les réponses de la personne sélectionnée représentent le ménage lorsqu'on utilise le poids des ménages.

Les utilisateurs familiers avec l'ESCC remarqueront que les ajustements de poids reliés aux poids des ménages sont similaires aux ajustements des autres poids au niveau de la personne qui étaient créés dans le passé. Une interview de l'ESCC peut être vue comme un processus en deux étapes. D'abord, l'intervieweur effectue le listage des membres du ménage (réponse de ménage). Ensuite, il/elle interviewe la personne sélectionnée dans le ménage (réponse de personne). Dans le calcul du poids des ménages, on utilise les réponses individuelles pour représenter les ménages. Cependant, les ajustements de non-réponse pour les deux étapes sont effectués lors de la production des poids des ménages, comme lors de la production des poids au niveau de la personne, parce que la non-réponse peut survenir à n'importe laquelle des étapes du processus de collecte. Les ajustements de non-réponse du ménage sont donc basés sur les caractéristiques de l'individu répondant. Étant donné que l'enquête est désignée pour recueillir les renseignements de l'individu, les caractéristiques de ce dernier peuvent avoir un effet sur la non-réponse du ménage.

2. Vue d'ensemble de la pondération

Pour que les estimations produites à partir de données d'enquête soient représentatives de la population couverte, et non pas seulement représentatif 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 ménages représentés par le répondant dans l'ensemble de la population.

Le ESCC de 2007 a eu 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 traite séparément la base aréolaire et la base téléphonique. Les poids résultant de ces deux bases sont ensuite combinés en un seul ensemble de poids lors d'une étape appelée « intégration ». Suite à quelques ajustements, ce poids intégré devient le poids final. Noter que dépendamment du besoin, une seule ou deux bases pouvaient être utilisées pour la sélection de l'échantillon dans une région sociosanitaire donnée. La stratégie de pondération s'occupe de cette particularité lors de l'étape d'intégration.

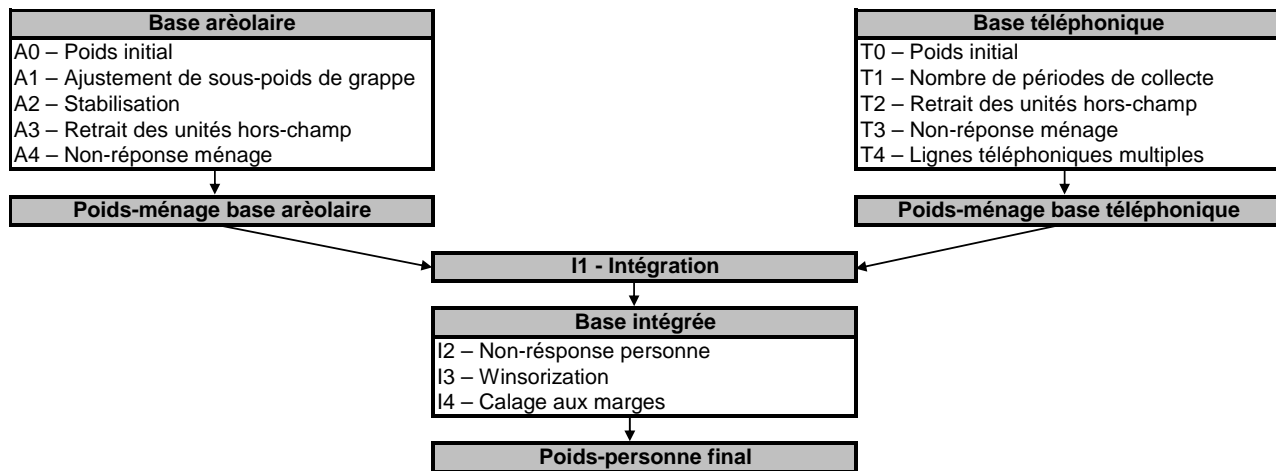
3. Pondération de l'échantillon

Tel que mentionné avant, les unités des bases aréolaire et téléphoniques sont traitées séparément jusqu'à l'étape d'intégration. Ces étapes de la pondération des ménages sont les mêmes que ceux du poids principal jusqu'à l'intégration des bases (inclusivement). Veuillez faire référence à la guide de l'utilisateur de l'ESCC de 2007 pour des renseignements plus complets concernant ces étapes. Les trois étapes finales de la pondération, c'est-à-dire l'ajustement pour la non réponse personnelle, la winsorization et la calage aux marges, qui sont expliquées dans les sous-sections 3.1 à 3.3.

Malgré que les deux bases aient été utilisées pour couvrir les trois territoires, certaines modifications ont dû être faits relativement à leur utilisation. Ces modifications ont affectés substantiellement la pondération de ces trois régions, et celles-ci sont rapportées dans la sous-section 3.4.

Le diagramme A présente un sommaire des différents ajustements qui font partie de la stratégie de production des poids des ménages dans l'ordre qu'ils sont appliqués. 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 les ajustements d'Intégration et ceux qui suivent.

Diagramme A : Sommaire de la stratégie de pondération (Poids des ménages)



3.1 Non-réponse personne (I2)

Dans le cadre de l'ESCC, une interview peut être vue comme un processus en deux étapes. Dans un premier temps, l'intervieweur obtient la liste complète des personnes vivant dans le ménage, puis par la suite, interviewe la personne sélectionnée dans le ménage. Dans certains cas, les intervieweurs ne réussissent qu'à compléter la première étape, soit parce qu'ils ne peuvent entrer en contact avec la personne sélectionnée, ou encore parce que la personne sélectionnée refuse d'être interviewée. Des 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 pour cette non-réponse. Tout comme pour la non-réponse à l'échelle du ménage (voir le Guide de l'utilisateur, section 8.2 – A4) l'ajustement est appliqué à l'intérieur des classes définies à partir des caractéristiques disponibles pour les répondants et les non-répondants. Toutes les caractéristiques recueillies lors du listage des membres du ménage, en plus des données géographiques étaient, en fait, disponibles pour créer ces classes. La méthode du score a été utilisée afin de définir les classes. Enfin, un facteur d'ajustement est calculé à l'intérieur de chaque classe de la façon suivante :

$$\frac{\text{Somme des poids } I1 \text{ pour toutes les personnes sélectionnées}}{\text{Somme des poids } I1 \text{ pour toutes les personnes sélectionnées répondantes}}$$

Le poids I1 a donc été multiplié par ce facteur d'ajustement pour produire le poids I2. Les personnes non-répondantes sont éliminées de la pondération à partir de ce point.

3.2 Winsorization (I3)

Noter que suite à la série d'ajustements appliqués sur les poids, il est possible que certaines unités se sont retrouvées avec des poids se démarquant des autres unités dans le même domaine d'intérêt au point même de devenir aberrants. Pour les poids des ménages, ces domaines incluent la province par la taille du ménage où la taille du ménage est définie par : ménage d'une personne, ménage de 2 personnes et ménage d'au moins 3 personnes. Certains répondants peuvent

effectivement représenter une partie anormalement élevée de leur croisement Province-taille du ménage et ainsi influencer fortement la variance. Afin d'éviter cette situation, le poids des unités aberrants qui représente une grande partie de leur domaine ont été ajustés à la baisse selon une méthode « winsorization ».

3.3 Calage aux marges (I4)

La dernière étape nécessaire afin d'obtenir le poids final des ménages de l'ESCC de 2007 était la calage aux marges. Le calage aux marges est appliqué en utilisant CALMAR afin d'assurer que la somme des poids finaux corresponde aux estimations du compte de ménages définies à l'échelle des provinces par taille du ménage. Ces groupes d'intérêts étaient définies par les tailles : ménage d'une personne, ménage de 2-personnes et ménage d'au moins 3 personnes. En même temps, les poids ont été ajustés de façon saisonnier afin d'assurer que chaque période de collection (période de deux mois) soit représenté de façon égale dans l'échantillon. En ce qui concerne la géographie, tout le calage aux marges a été effectué au niveau provincial seulement.

Les estimations du compte de ménages ont été basées sur le plus récent recensement. Les moyennes de ces estimations mensuelles ont été utilisées pour la calage aux marges pour chaque croisement Province-Taille du ménage, post-strate, à l'intérieur d'une période de collection. Le poids I3 a ainsi été ajusté pour obtenir le poids final I4, à l'aide du facteur d'ajustement I4 défini comme suit :

$$\frac{\text{Estimation de population pour le croisement Province - Taille du ménage}}{\text{Somme des poids I3 pour le croisement Province - Taille du ménage}}$$

Par conséquent, le poids I4 correspond au *poids de ménage final de l'ESCC de 2007* que l'on retrouve dans le fichier de poids-ménage portant le nom de variable WTS_MHH pour le poids maître et WTS_SHH pour le poids partagé.

3.4 Particularités de la pondération pour les trois territoires

Pour l'ESCC de 2007, le plan d'échantillonnage utilisé pour les trois territoires était quelque peu différent de celui utilisé dans les 10 provinces. La stratégie de pondération a donc dû être adaptée pour répondre à ces différences. Cette section résume les changements apportés aux étapes de pondération pour les territoires.

D'abord pour la base aréolaire, une étape additionnelle de sélection a été ajoutée pour les territoires. Chaque territoire était initialement stratifié selon des regroupements de communautés à l'intérieur desquels, on a sélectionné aléatoirement une communauté à l'intérieur de chaque groupe. Noter que les capitales de chaque territoire formaient une strate à elles seules, et étaient donc toutes trois sélectionnées automatiquement lors de ce premier stage. Cette particularité 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) a été appliquée aux unités de la base aréolaire. Les classes d'ajustement pour les ménages et personne ont été construites de la même manière que les provinces à l'aide du même ensemble de variables disponibles.

Pour ce qui est de la pondération des unités de la base téléphonique, mentionnons tout d'abord que seule la base CA a été utilisée pour les territoires, et ce, uniquement à l'intérieur des capitales du Yukon et des Territoires du Nord-Ouest. Tous les ajustements téléphoniques standards ont été appliqués.

Les deux ensembles de poids (aréolaire et téléphonique) ont ensuite été intégrés, puis ajustés pour la nonréponse personnelle, winsorization et enfin post-stratifiés de façon semblable à ce qui a été fait pour les provinces, à l'exception de trois détails. D'abord, l'intégration a été appliquée uniquement pour les unités situées dans les capitales du Yukon et des Territoires du Nord-Ouest; les autres communautés ayant été couvertes uniquement par la base aréolaire. De plus, pour le Nunavut, les comptes de la population des ménages utilisés pour la calage aux marges représentent les dix plus grandes communautés (et 73% des ménages seulement) étant donné la sous couverture de la base aréolaire (voir la section 5.4.1 du Guide de l'utilisateur pour plus de détails). Enfin, deux périodes de six-mois ont été utilisés dans le post-stratification pour l'effet saisonnier (au lieu de six périodes de deux mois) dû aux différences dans les stratégies de collection.