

Enquête sur la santé dans les collectivités canadiennes (ESCC)

L'objectif principal de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) est la collecte de données dans le domaine de la santé, à des niveaux géographiques infraprovinciaux (régions sociosanitaires ou régions sociosanitaires regroupées).

Information détaillée pour 2005 (cycle 3.1)

Date de la parution – 21 décembre 2005 (cette diffusion de données couvre les données collectées de janvier à juin 2005)

Aperçu de l'enquête

En 1991, le Groupe de travail national sur l'information en matière de santé a relevé plusieurs 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 canadienne. 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 national de l'information sur la santé.

Notre système d'information sur la santé devrait nous fournir l'information pour répondre aux questions cruciales ci-dessous:

1. À quel point le système de santé est-il sain?
2. À quel point les Canadiens sont-ils en santé?

Afin de répondre à ces questions l'enquête sur la santé dans les collectivités canadiennes (ESCC) a été créée. Les objectifs principaux de l'ESCC sont les suivants :

- fournir des estimations transversales à jour et fiables des déterminants de la santé, de l'état de santé et de l'utilisation des services de santé à travers le Canada;
- recueillir des données au niveau infraprovincial;
- créer un instrument d'enquête souple permettant :
 - de combler des lacunes statistiques particulières au niveau de la région sociosanitaire,
 - d'élaborer un contenu d'enquête thématique en vue de recueillir des données importantes;
 - de répondre aux nouvelles questions ayant trait à la santé et aux services de santé à mesure qu'elles surviennent.

L'ESCC, en tant que composante importante du Programme des enquêtes sur la santé, permet de combler des besoins d'information accrus en matière de santé. Il s'agit de :

- faciliter l'élaboration de politiques gouvernementales;
- fournir des données permettant de réaliser des études analytiques qui aideront à comprendre les déterminants de la santé;

- recueillir des données sur les corrélations entre la santé et les facteurs économiques, sociaux, démographiques, professionnels et environnementaux;
- permettre de mieux comprendre la relation entre l'état de santé et l'utilisation des services de santé.

Au Canada, l'utilisation principale de ces données est la surveillance de la santé, comme la prévalence de maladies et autres formes de recherche en santé. Les données sont utilisées de manière intensive par la communauté des chercheurs et autres professionnels de la santé. L'aspect unique de l'ESCC est la nature régionale de son contenu et de la mise en œuvre de l'enquête. Ces aspects permettent de réaliser des analyses au niveau régional, à travers le Canada. Les ministères fédéraux et provinciaux de la santé et des ressources humaines, les organismes de services sociaux et autres agences gouvernementales utilisent l'information recueillie des répondants afin de planifier, implémenter et évaluer des programmes pour améliorer la santé des Canadiens et l'efficacité des services de santé. Les organisations à but non lucratif et les chercheurs du milieu académique utilisent l'information afin de faire avancer la recherche et améliorer la santé. Les médias utilisent les résultats des enquêtes afin de sensibiliser la population aux questions relatives à la santé qui concernent tous les Canadiens.

Période de référence

Varie selon la question (par exemple: « durant les 12 derniers mois », « durant les 6 derniers mois » », « durant la dernière semaine », etc.)

Notes aux utilisateurs relativement à cette diffusion

VEUILLEZ NOTER : Cette diffusion de données couvre les données collectées au cours des 6 premiers mois (janvier à juin 2005) du cycle 3.1 de l'ESCC. À ce stade, l'enquête avait recueilli des renseignements sur environ 68 000 personnes âgées de 12 ans et plus.

Seulement une partie de l'information collectée à partir du questionnaire de l'ESCC 3.1 a été traitée et finalisée pour cette diffusion. L'ensemble des données couvrant la période entière de collecte de 12 mois (janvier à décembre 2005) seront diffusées au cours de l'été 2006.

Population cible

Les données de l'ESCC sont recueillies auprès des personnes de 12 ans et plus vivant dans des logements privés dans les dix provinces et les trois territoires. Sont exclus de la base de sondage les habitants des réserves indiennes et des terres de la Couronne, les personnes vivant en institution, les membres à temps plein des Forces armées canadiennes et les habitants de certaines régions éloignées. L'ESCC couvre environ 98 % de la population canadienne âgée de 12 ans et plus.

Élaboration de l'instrument

Les questionnaires de chacun des cycles de l'ESCC ont tous été conçus en collaboration avec des spécialistes provenant de Statistique Canada, d'autres ministères et/ou du milieu académique. Les questions de l'ESCC ont été conçues pour être posées selon la méthode d'interview assistée par ordinateur (IAO), ce qui veut dire que l'ordre logique des questions a été programmé à mesure que celles-ci ont été élaborées. On a notamment spécifié le genre de réponse voulue, les valeurs minimales et maximales, les vérifications en ligne liées aux questions et la suite à donner en cas de non-réponse.

L'IAO permet de contrôler l'interview d'après les déclarations du répondant. Lorsqu'une réponse invalide est saisie un message s'affiche à l'écran, ce qui permet au répondant et/ou à l'intervieweur d'obtenir un retour d'information immédiat pour corriger les incohérences. L'ajout automatique de la période de référence d'après la date du jour est un autre exemple d'amélioration. Il est possible de pré-insérer du texte ou des données selon les renseignements recueillis durant l'interview, ce qui permet à l'intervieweur de continuer son travail sans devoir revenir en arrière pour trouver des réponses antérieures. Cette possibilité permet notamment d'insérer le nom du répondant ou de faire les accords en genre à l'intérieur même des questions. Il est également possible de programmer des intervalles ou des réponses acceptables d'après les données recueillies durant l'interview. En d'autres termes, le questionnaire peut être adapté au répondant selon les données recueillies au moment de l'interview ou durant une interview précédente.

Un essai sur le terrain a été effectué avant le cycle 3.1. Les bureaux régionaux de Statistique Canada ont participé à cet essai. Les interviews ont été menées par des intervieweurs expérimentés de l'Enquête sur la population active. Cet essai visait principalement à observer les réactions des répondants, à estimer la durée des diverses sections du questionnaire, à déterminer les taux de réponse et à évaluer les questions avec rétroaction. Les opérations et les procédures sur le terrain, la formation des intervieweurs et l'application informatique de collecte de donnée ont aussi été soumises à des essais. En plus de l'essai sur le terrain, l'application informatique de collecte des données a été soumise à des essais à grande échelle à l'interne pour repérer toute erreur dans le cheminement du programme ou dans le texte. La mise à l'essai de l'application s'est poursuivie de façon continue jusqu'au début de l'enquête principale.

Plan d'échantillonnage

Afin de produire des estimations fiables pour les 122 régions sociosanitaires (RSS) et compte tenu du budget accordé pour le cycle 3.1 de l'ESCC, il a été établi que cette composante de l'enquête devrait être réalisée auprès d'un échantillon de 130 000 personnes. Les 3 territoires ont été exclus de cette diffusion de 6 mois de données en raison de leurs faibles échantillons. Un total d'approximativement 68 000 répondants couvrant 119 RSS sont inclus dans le fichier de 6 mois. La stratégie de répartition de l'échantillon, qui comporte trois étapes, accorde une importance plus ou moins égale aux RSS et aux provinces. Lors des deux premières étapes, l'échantillon a été réparti entre les provinces en fonction de leur compte de population et du nombre de RSS qu'elles

contiennent. 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.

L'ESCC a utilisé trois bases de sondage pour sélectionner les échantillons de ménages: 50 % de l'échantillon des ménages est venu de la base aréolaire, 49 % est venu d'une base liste de numéros de téléphone et le 1 % restant est venu d'une base d'échantillonnage à composition aléatoire (CA). Pour la plupart des régions sociosanitaires, 50 % de l'échantillon a été sélectionné à partir d'une base aréolaire et 50 % d'une base liste de numéros de téléphone. Dans deux régions sociosanitaires (Nord du Québec et Nord de la Saskatchewan), seule la composition aléatoire a été utilisée. Au Nunavut, seule la base aréolaire a été utilisée. Au Yukon et aux Territoires du Nord-Ouest, la plupart de l'échantillon est venu de la base aréolaire mais un petit échantillon de composition aléatoire a aussi été sélectionné à Whitehorse et à Yellowknife.

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'échantillonnage. À la première étape, on a formé des strates homogènes et sélectionné un échantillon indépendant de grappes, dans chaque strate. À la deuxième étape, on a dressé la liste des logements pour chaque grappe, puis on a sélectionné les logements, ou les ménages, d'après les listes.

Pour les besoins du plan d'échantillonnage, chaque province est répartie en trois catégories de région, à savoir 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. Dans certains centres urbains, des strates distinctes sont créées pour les immeubles à appartements ou les secteurs de dénombrement (SD) pour lesquels le revenu moyen du ménage est élevé. Dans chaque strate, on sélectionne six grappes ou immeubles résidentiels (pouvant compter de 12 à 18 appartements) par une méthode d'échantillonnage aléatoire avec probabilité proportionnelle à la taille (PPT), cette dernière correspondant au nombre de ménages. Le nombre 6 est utilisé pour l'ensemble du plan d'échantillonnage afin de permettre le renouvellement mensuel d'un sixième de l'échantillon de l'EPA.

Les autres villes et régions rurales de chaque province sont stratifiées, en premier lieu, en fonction de données géographiques, puis selon les caractéristiques socioéconomiques. Dans la plupart des strates, on sélectionne six grappes (habituellement des SD) par la méthode PPT. Pour les strates où la densité de population est faible, on suit un plan en trois étapes en vertu duquel on sélectionne deux ou trois unités primaires d'échantillonnage (UPE), qui correspondent normalement à des groupes de SD, puis on les répartit en grappes dont six sont sélectionnées pour faire partie de l'échantillon. On obtient l'échantillon final par échantillonnage systématique des logements.

Sources des données

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

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

Les intervieweurs IPAO travaillent séparément depuis leur domicile à l'aide d'un ordinateur portatif et sont supervisés à distance par des intervieweurs principaux. Les interviews complétées sont envoyées quotidiennement de leur domicile au bureau central de Statistique Canada, et ce, par transmission téléphonique protégée. Les intervieweurs ITAO travaillent dans des bureaux centralisés sous la supervision d'un intervieweur principal. Il incombe au surveillant de projet du bureau régional, à l'intervieweur principal et à l'équipe d'assistance technique de transmettre les dossiers de chacun des cinq bureaux ITAO au bureau central.

On a pu compter pour la collecte par ITAO, sur un ordonnanceur d'appels automatique, c'est-à-dire un système central qui optimise l'agencement des rappels et l'ordonnement des rendez-vous.

Détection des erreurs

La vérification des données a été exécutée en grande partie par l'application d'interview assistée par ordinateur (IAO) durant la collecte des données. Il était impossible que les intervieweurs saisissent des valeurs situées en dehors de la fourchette permise et qu'ils commettent des erreurs d'enchaînement, car celui-ci était contrôlé par le système programmé de sauts de question (commandes « passez à ») de l'IAO. Par exemple, l'IAO fait en sorte que les questions qui ne s'appliquent pas au répondant ne sont pas posées. Dans le cas de certains types de déclarations non cohérentes ou inhabituelles, l'application a affiché des messages d'avertissement, mais aucune mesure corrective n'a été prise au moment de l'interview. Des vérifications élaborées pour ce genre de déclaration ont été effectuées, au besoin, après la collecte des données, au Bureau central. En général, on a corrigé les incohérences en attribuant la valeur « Non déclaré » à l'une des variables en cause ou aux deux.

Imputation

Il n'y a eu aucune imputation.

Estimation

Le principe sur lequel s'appuie l'estimation dans le cas d'un échantillon probabiliste tel que celui de l'ESCC veut que, outre elle-même, chaque personne faisant partie de l'échantillon en « représente » plusieurs autres qui ne font pas partie de l'échantillon. Par exemple, dans un échantillon aléatoire simple au 1/50 (2 %) de la population, chaque personne représente 50 membres de la population. Conformément à la terminologie utilisée ici, chaque personne a un coefficient de pondération (ou poids) de 50. La phase de pondération est l'étape où l'on calcule ce poids d'échantillonnage pour chaque personne. Le poids figure dans le fichier de microdonnées et doit être utilisé pour produire des estimations représentatives d'après les données de l'enquête. Par exemple, si l'on veut estimer le nombre de personnes qui ont vécu un épisode dépressif majeur, on choisit les enregistrements se rapportant aux personnes faisant partie de l'échantillon qui présentent cette caractéristique et l'on additionne les poids inscrits sur ces enregistrements. Afin que les estimations produites à partir des données d'enquête soient représentatives de la population couverte et non seulement de l'échantillon lui-même, un utilisateur doit incorporer les poids dans les calculs. Un poids est donné à chaque personne incluse dans l'échantillon final, c'est-à-dire l'échantillon des personnes qui ont répondu à l'enquête. Ce poids correspond au nombre de personnes représentées par le répondant pour la population entière.

Afin de déterminer la qualité de l'estimation et de calculer le coefficient de variation (CV), il est nécessaire de calculer l'écart type. Les intervalles de confiance nécessitent également le calcul de l'écart type de l'estimation. Pour l'ESCC, on utilise un plan de sondage à plusieurs degrés, ce qui signifie qu'il n'y a pas de formule simple pour calculer les estimations de la variance. Par conséquent, il a fallu appliquer une méthode approximative. On applique la méthode bootstrap parce que l'information sur le plan d'échantillonnage doit être prise en compte lors du calcul des estimations de la variance. La méthode bootstrap permet d'effectuer ce calcul et, conjointement avec le programme Bootvar décrit à la sous-section suivante, elle demeure une méthode relativement facile à utiliser. La méthode de ré-échantillonnage bootstrap utilisée pour l'ESCC sous-tend la sélection d'échantillons aléatoires simples, appelés les échantillons répétés, et le calcul de l'écart entre les estimations d'un échantillon répété à l'autre. Dans chaque strate, on sélectionne un échantillon aléatoire simple de $(n-1)$ des n grappes avec remise pour former un échantillon répété. Il est à noter que puisque la sélection se fait avec remplacement, une grappe peut être choisie plus d'une fois. Pour chaque échantillon répété, on recalcule le poids de sondage de chaque enregistrement dans les grappes $(n-1)$ sélectionnées. Ces poids sont par la suite post-stratifiés en fonction de l'information démographique de la même façon que les pondérations du plan de sondage, ce qui permet d'obtenir les poids bootstrap finaux. Le processus complet (sélectionner les échantillons aléatoires simples, recalculer et post-stratifier les poids de chaque strate) est répété B fois,

B prenant une grande valeur. En général, on utilise $B=500$ pour l'ESCC afin de produire 500 poids bootstrap. Pour obtenir l'estimateur bootstrap de la variance, on doit calculer l'estimation ponctuelle de chacun des B échantillons. L'écart type de ces estimations représente l'estimateur bootstrap de la variance. Statistique Canada a élaboré un programme qui peut effectuer tous ces calculs pour l'utilisateur : le programme Bootvar.

Évaluation de la qualité

La conception de l'enquête exerce une influence profonde sur ses objectifs, énumérés sous la rubrique « Description de l'enquête ». Pour atteindre ces objectifs, un Comité directeur et un Conseil consultatif composés de responsables des ministères provinciaux de la Santé et de Santé Canada ont fixé les concepts et les centres d'intérêt. Ils ont réuni des groupes d'experts pour les conseiller sur les mesures à prendre afin d'obtenir les résultats envisagés et pour leur recommander des instruments de collecte et des indices éprouvés. Les données qui en résultent sont reconnues comme mesures valables de concepts contemporains tels que la dépression, la limitation de l'activité, les problèmes de poids et la douleur chronique.

Les bases de sondage servant à tirer l'échantillon, l'Enquête sur la population active, ITAO et les listes de numéros de téléphone, ont été combinées à des méthodologies de plan d'échantillonnage vérifiées, utilisées à plusieurs reprises et ayant fait la preuve de la qualité de leurs résultats. Le grand échantillon de chaque province et territoire contribue à la garantie de résultats exacts et significatifs.

Un taux de réponse élevé est essentiel à la qualité des données. On a pris des mesures pour réduire au minimum les erreurs non dues à l'échantillonnage. Pour réduire le nombre de cas de non-réponse, les intervieweurs de Statistique Canada reçoivent une formation intensive, disposent de manuels détaillés et travaillent sous la direction de superviseurs. L'ampleur de la non-réponse varie de non-réponse partielle (le fait de ne pas répondre à une ou plusieurs questions) à une non-réponse totale. Dans le cas de l'ESCC, il n'y a presque pas eu de non-réponse partielle car une fois le questionnaire débuté les répondants avaient tendance à le terminer. Dans la plupart des cas, la non-réponse partielle à l'enquête était attribuable au fait que le répondant ne comprenait pas une question ou l'interprétait mal, refusait d'y répondre, ne se souvenait pas de l'information demandée ou n'était pas en mesure de fournir de l'information sur lui-même ou sur la personne pour laquelle il fournissait une réponse par procuration. Il y a eu non-réponse totale lorsqu'il était impossible de dépister le répondant, qu'aucun membre du ménage ne pouvait fournir les renseignements voulus ou que le répondant refusait de participer à l'enquête. On a traité les cas de non-réponse en corrigeant les poids des personnes qui ont répondu à l'enquête afin de compenser pour ceux qui n'ont pas répondu.

Contrôle de la divulgation

La Loi interdit à Statistique Canada de rendre publique toute donnée susceptible de révéler de l'information obtenue en vertu de la Loi sur la statistique et se rapportant à

toute personne, entreprise ou organisation reconnaissable sans que cette personne, entreprise ou organisation le sache ou y consente par écrit. Diverses règles de confidentialité s'appliquent à toutes les données diffusées ou publiées afin d'empêcher la publication ou la divulgation de toute information jugée confidentielle. Au besoin, des données sont supprimées pour empêcher la divulgation directe ou par recoupement de données reconnaissables.

On planifie la diffusion de Fichiers de microdonnées à grande diffusion (FMGD) à partir des données de l'ensemble des 12 mois de collecte du cycle 3.1 de l'ESCC à l'été 2006. Les FMGD diffèrent des fichiers maîtres détenus par Statistique Canada en de nombreux points. Ces différences sont le résultat d'actions prises afin de protéger l'anonymat des répondants individuels aux enquêtes. Premièrement, seules les données transversales sont disponibles sur de tels fichiers, car l'information longitudinale permet d'identifier des répondants. Aussi, quelques variables sensibles ont été regroupées, maximisées ou complètement enlevées des fichiers. Les utilisateurs devant avoir accès à l'information exclue des fichiers de microdonnées peuvent acheter des totalisations spéciales, ou accéder aux fichiers maîtres par l'entremise du programme des Centres de données de recherche ou par le programme de télé accès. Les résultats sont vérifiés pour respecter la confidentialité avant d'être fournis aux utilisateurs.

Avant de diffuser et/ou de publier des estimations tirées de ces fichiers de microdonnées, les utilisateurs doivent d'abord déterminer le nombre de répondants qui ont fourni les données entrant dans le calcul de l'estimation. 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 estimations pondérées basées sur des échantillons composés de 30 répondants ou plus, les utilisateurs doivent calculer le coefficient de variation de l'estimation arrondie et suivre les lignes directrices ci-après.

Les estimations qui figurent dans le corps d'un tableau statistique doivent être arrondies à la centaine près selon la méthode d'arrondissement classique. Si le premier ou le seul chiffre à supprimer se situe entre zéro et quatre, le dernier chiffre qu'on garde ne change pas. Si le premier ou le seul chiffre à supprimer se situe entre cinq et neuf, on augmente d'une unité la valeur du dernier chiffre à garder. Les totaux partiels marginaux et les totaux marginaux des tableaux statistiques doivent être calculés d'après leurs éléments correspondants non arrondis, puis arrondis à leur tour à la centaine près selon la méthode d'arrondissement classique. Les moyennes, les proportions, les taux et les pourcentages doivent être calculés à partir d'éléments non arrondis (par exemple les numérateurs ou les dénominateurs), puis arrondis à une décimale près selon la méthode d'arrondissement classique. Pour arrondir à un seul chiffre selon cette méthode classique, si le dernier ou le seul chiffre à supprimer se situe entre zéro et quatre, le dernier chiffre à garder ne change pas. Si le premier ou le seul chiffre à supprimer se situe entre cinq et neuf, on augmente d'une unité le dernier chiffre à garder. Les sommes et les différences d'agrégat (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 première décimale près) selon la méthode d'arrondissement classique. On ne doit en aucun cas publier ni diffuser d'une autre façon des estimations non arrondies. Les estimations non arrondies donnent l'impression d'être beaucoup plus précises qu'elles ne le sont en réalité.