

Mesures de qualité

Enquête nationale sur la santé de la population - Volet ménages

(Numéro de l'Enquête 3236)

Couverture

La population cible de la composante des ménages de l'ENSP comprend les résidents des ménages de toutes les provinces du Canada en 1994-1995, sauf les habitants des réserves indiennes et des terres de la Couronne, les résidents des établissements de santé, les membres à temps plein des Forces canadiennes et les habitants de certaines régions éloignées de l'Ontario et du Québec.

Les erreurs de couverture ont lieu à plusieurs étapes de l'enquête, durant la conception de la base de sondage, la définition des unités d'échantillonnage, ainsi que la collecte et le traitement des données. L'indicateur utilisé pour évaluer l'erreur de couverture est le *taux de glissement*, qui représente l'écart entre les estimations de population (*d'après l'enquête*) (*sans utilisation de données externes, c'est-à-dire fondées sur les poids stratifiés a priori et a posteriori*) et les estimations démographiques les plus récentes d'après le recensement. On exprime l'écart en pourcentage de l'estimation d'après le recensement. Comme la plupart des enquêtes, l'ENSP produit un certain sous-dénombrement qui se manifeste par un taux de glissement positif (*environ 10 % pour le panel longitudinal sélectionné en 1994-1995*).

L'étude du glissement est décrite dans un document de travail sur la méthodologie qui sera disponible à l'automne 2003.

Pour réduire l'effet de l'erreur de couverture, les poids d'échantillonnage sont rajustés durant leur calcul en se basant sur les estimations de population fournies pour la période de référence de l'enquête.

Réponse totale

Nota : le calcul du taux de réponse au cycle 1 n'est pas le même que celui des taux de réponse aux trois autres cycles. Le taux de réponse au cycle 1 est fondé sur les 10 095 personnes comprises dans le champ d'observation de l'enquête sélectionnées pour former le panel, tandis que les taux de réponse des cycles subséquents sont fondés sur les 17 276 personnes qui constituent effectivement le panel longitudinal.

<u>Cycle</u>	<u>Taux de réponse</u>
Cycle 1	3,6 %
Cycle 2	92,8 %
Cycle 3	88,2 %
Cycle 4	84,8 %

L'érosion est une réduction de l'effectif de l'échantillon due à la non-réponse, c'est-à-dire les personnes qui refusent de participer, celles qu'on ne peut rejoindre, celles qu'on ne peut dépister lors du suivi, etc. Le taux d'érosion cumulatif est présenté pour chaque cycle. Chaque taux d'érosion est calculé en utilisant le nombre de personnes comprises dans le sous-ensemble de répondants à réponse complète, c'est-à-dire ceux qui ont répondu au questionnaire lors de chaque cycle. La cause principale d'érosion est le nombre croissant de répondants qui refusent de continuer de participer à l'enquête.

Cycle	Taux d'érosion
Cycle 2	9,3 %
Cycle 3	15,4 %
Cycle 4	21,4 %

Afin d'essayer de réduire l'érosion due aux refus, on a décidé de remanier le matériel distribué aux répondants pour le cycle 6 de l'ENSP. En outre, on explore diverses stratégies en vue de convaincre les répondants de s'associer à SC et de continuer à répondre à l'ENSP.

Méthodes utilisées pour corriger pour la non-réponse totale : La stratégie de pondération élaborée pour chaque cycle de l'ENSP comprend divers ajustements, y compris ceux pour la non-réponse au niveau du ménage et au niveau de la personne. Pour tenir compte de la non-réponse totale, on rajuste le poids appliqué aux personnes qui ont répondu à l'enquête, de façon à tenir compte de celles qui n'ont pas répondu. Cette correction est faite par catégorie de non-réponse, c'est-à-dire les groupes d'unités présentant la même propension à répondre à l'enquête, afin de minimiser le biais dû à la non-réponse.

Méthodes utilisées pour augmenter la réponse totale : Diverses stratégies sont suivies durant la collecte des données pour améliorer les taux de réponse, comme la formation des intervieweurs, l'envoi de lettres d'introduction, la réalisation des interviews dans d'autres langues que le français ou l'anglais, les suivis auprès des non-répondants, la surveillance du taux de réponse et le transfert des charges de travail à d'autres bureaux.

Taille de l'échantillon et erreur d'échantillonnage

L'échantillon longitudinal comprend les 17 276 (*sur les 20 095*) personnes qui ont été sélectionnées au cycle 1 et qui ont répondu au moins à la composante générale du questionnaire au cycle 1. Ce panel a été interrogé lors des cycles 2, 3, 4 et 5, et sera interrogé lors de futurs cycles de l'ENSP. Les échantillons supplémentaires ajoutés aux cycles 1, 2 et 3 aux fins d'analyses transversales ne font pas partie de l'échantillon longitudinal. Celui-ci n'est pas renouvelé au cours du temps. Aucun membre du panel n'a été, ni se sera considéré comme étant hors du champ de l'enquête. La taille de l'échantillon longitudinal reste la même, c'est-à-dire 17 276, pour tous les cycles.

Méthodes utilisées pour estimer l'erreur d'échantillonnage : La variance et les coefficients de variation (c.v.) exacts des estimations produites d'après ces fichiers de données sont calculés par la méthode du bootstrap. Pour chaque cycle, les partenaires

avec lesquels des données sont échangées et les Centres de données de recherche reçoivent un fichier qui contient les poids bootstrap, un programme pour le calcul de la variance et des c.v. pour un certain nombre de statistiques et une documentation complète. Les tableaux de données de l'ENSP qui figurent dans CANSIM incluent les intervalles de confiance ou un indicateur type de la précision des estimations fondé sur le c.v. et les lignes directrices de Statistique Canada. Par exemple, les estimations dont le c.v. est supérieur à 33,3 \$ sont supprimées des tableaux et les cellules correspondantes sont marquées d'un « F ».

Réponse partielle

La non-réponse partielle au questionnaire du cycle 4 est d'environ 0,09 %. Un taux de non-réponse plus élevé est observé pour quelques variables, comme le revenu du ménage (3,8 %) et le revenu personnel (1,5 %), ainsi que certaines variables de stress (environ 0,6 %).

Diverses mesures sont prises pour réduire la non-réponse partielle au minimum. Les intervieweurs reçoivent une formation concernant l'objectif et les concepts de toutes les questions. Les données de l'ENSP sont recueillies au moyen d'une application d'interview assistée par ordinateur (IAO) qui assure que toutes les questions appropriées, et uniquement ces questions, soient posées.

Méthodes d'ajustement pour la non-réponse partielle : Au début de l'enquête, il a été décidé, en collaboration avec les analystes de SC, qu'on ne ferait aucun ajustement pour la non-réponse partielle.

Dépistage et correction des erreurs

La plupart des vérifications sont réalisées au moment de l'interview au moyen de l'application d'IAO. Certains types de réponses incohérentes ou inhabituelles sont vérifiées au Bureau central après la collecte des données. Habituellement, on corrige les incohérences en fixant la réponse à une question à « Non déclaré ». Font exception les vérifications des liens, pour lesquels les incohérences décelées sont corrigées manuellement. Aucune imputation n'est faite sur les données de la composante des ménages de l'ENSP.

Autres questions concernant l'exactitude des données

Une documentation détaillée est produite pour chaque cycle de la composante des ménages de l'ENSP afin de fournir aux utilisateurs des données, tous les renseignements contextuels pertinents sur l'enquête (*contexte, méthodologie, qualité des données, dictionnaire de données, spécification des variables dérivées, clichés d'enregistrement, etc.*). En outre, des renseignements comparables sont chargés dans la Base de métadonnées intégrée (BMDI).

Un article où sont décrits certains indicateurs de la qualité des données de l'ENSP, comme le nombre de tentatives de prise de contact, la durée de l'interview, la non-réponse totale/partielle (*refus/ne sait pas*) et les rejets à la vérification, a été rédigé par la Division des méthodes auprès des ménages (*voir Bibliographie*). L'article fournit aussi des renseignements sur le dépistage et la déclaration par procuration. D'autres aspects de la qualité des données, comme les taux de réponse et les taux d'érosion de l'échantillon sont décrits dans la documentation sur les fichiers longitudinaux du cycle 4.

D'autres aspects de la qualité des données, comme les profils et les caractéristiques de certains membres du panel, présentent aussi un intérêt. Certaines études internes portant sur la qualité des données ont été réalisées. Elles incluent les profils des non-répondants, des personnes qui refusent de participer, des personnes qu'il est impossible de dépister lors des suivis, des répondants par procuration, des personnes qui refusent d'autoriser le partage des renseignements qu'elles fournissent et celles qui refusent que les renseignements qu'elles fournissent soient couplés aux données d'autres fichiers. Bien que certains de ces profils, comme celui des non-répondants au moment de calculer les poids utilisés pour corriger pour la non-réponse, ont déjà été étudiés, mais des études plus formelles seront réalisées dans l'avenir.

Bibliographie

Cycle 4 (2000-2001) Household Component Longitudinal Documentation, Division de la statistique de la santé, Statistique Canada, non publié, mai 2002.

NPHS Data Quality : Exploring Non-sampling Errors, document de travail, Direction de la méthodologie, HSMD-2003-004E, Sandra Tolusso et François Brisebois, Division des méthodes d'enquête auprès des ménages, Statistique Canada, juillet 2003