

Article

Symposium 2008 :
Collecte des données : défis, réalisations et nouvelles orientations

Évaluation de la couverture des données de l'Enquête sur la santé dans les collectivités canadiennes couplées à celles des dossiers d'hospitalisation

par Michelle Rotermann

2009



Évaluation de la couverture des données de l'Enquête sur la santé dans les collectivités canadiennes couplées à celles des dossiers d'hospitalisation

Michelle Rotermann¹

Résumé

Contexte : L'évaluation de la couverture du fichier obtenu par couplage de données administratives recueillies régulièrement par les hôpitaux et de données d'enquête est une importante étape préalable aux analyses fondées sur ce fichier couplé. Données et méthodes : Afin d'évaluer la couverture du couplage des données du cycle 1.1 de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) aux données sur les hospitalisations (Information-santé orientée vers la personne, ou ISOP), le nombre de personnes admises à l'hôpital selon l'ISOP a été comparé à l'estimation pondérée obtenue pour les répondants à l'ESCC dont les données ont pu être couplées à celles de l'ISOP. Les écarts entre le chiffre selon l'ISOP et les estimations d'après les données couplées et pondérées de l'ESCC révèlent un échec du couplage ou un sous-dénombrement. Résultats : Selon l'ISOP, 1 572 343 personnes (hors Québec) âgées de 12 ans et plus ont été hospitalisées de septembre 2000 à la fin de novembre 2001. Les estimations fondées sur les données couplées de l'ESCC, rajustées pour tenir compte du consentement au couplage et de la vraisemblance du numéro d'assurance-maladie, sont inférieures de 7,7 % à ce chiffre. Les taux de couverture des hommes et des femmes sont comparables. Les taux provinciaux ne diffèrent pas de ceux observés pour le reste du Canada, mais des différences se dégagent pour les territoires. Les taux de couverture sont significativement plus faibles pour les personnes de 75 ans et plus que pour celles de 12 à 74 ans.

Mots clés : Couverture, enquêtes sur la santé, dossiers d'hospitalisation, couplage probabiliste.

1. Introduction

Dans les études sur la santé, le couplage d'enregistrements est utilisé afin d'obtenir des renseignements plus complets, de combler des lacunes dans les ensembles de données existants ou d'améliorer la qualité des données (Fair, 2004, Fair et Whitridge, 1997). La présente étude a été motivée par la nécessité d'évaluer la couverture du couplage des enregistrements de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) à ceux de l'Information-santé orientée vers la personne (ISOP), *une base de données administrative provenant des dossiers d'hospitalisation*. Une première étude du taux de couplage des enregistrements de l'ESCC et de l'ISOP a consisté à estimer la proportion de participants à l'ESCC qui avaient été hospitalisés entre 1994-1995 et 2004-2005, mais il reste à évaluer la couverture du fichier couplé (Nadeau et coll., 2006). Cette évaluation de la couverture est indispensable si le fichier couplé doit servir à la recherche épidémiologique. Il importe de déterminer si les conclusions seront biaisées, c'est-à-dire si les données sur les personnes possédant certaines caractéristiques sont plus susceptibles d'être couplées que celles sur les personnes qui ne les possèdent pas.

L'ISOP et l'ESCC sont des sources de données complémentaires. L'ISOP ne renferme pas de renseignements sur les déterminants non médicaux de la santé, comme les facteurs socioéconomiques et les facteurs liés au mode de vie. Ainsi, les dossiers d'hospitalisation ne renseignent pas sur l'usage du tabac ni l'indice de masse corporelle (IMC), deux facteurs de risque importants. En revanche, l'ESCC est riche en information sur l'état de santé et les déterminants de la santé, mais pauvre en détails nécessaires à l'étude de l'hospitalisation. La combinaison des données de l'ISOP et de l'ESCC réduit bon nombre des limites de chaque source et permet donc de mieux comprendre les raisons pour lesquelles les Canadiens entrent en contact avec le système de soins de santé et quels sont leurs résultats lorsqu'ils se trouvent dans le système.

¹ Michelle Rotermann, Division de l'information et de la recherche sur la santé, Statistique Canada, 100 promenade Tunney's Pasture, K1A 0T6, Ottawa, Canada (Michelle.Rotermann@statcan.gc.ca).

Les deux grands objectifs de la présente étude étaient de 1) calculer le taux de couverture du fichier de données couplées de l'ESCC et de l'ISOP et 2) dégager les caractéristiques des participants au cycle 1.1 de l'ESCC qui étaient moins susceptibles que les autres de figurer dans le fichier couplé.

2. Données et méthodes

2.1 Sources des données

2.1.1 Enquête sur la santé dans les collectivités canadiennes

L'Enquête sur la santé dans les collectivités canadiennes (ESCC) est une enquête transversale conçue pour recueillir des renseignements sur l'état de santé, l'utilisation des services de santé et les déterminants de la santé. Elle couvre la population à domicile de 12 ans et plus des provinces et des territoires, sauf les membres réguliers des Forces canadiennes et les personnes vivant en établissement, dans les réserves indiennes, dans d'autres établissements autochtones et en certaines régions éloignées. Le taux de couverture est de 98 % dans les provinces, de 97 % dans les Territoires du Nord-Ouest, de 90 % au Yukon et de 71 % au Nunavut. Les données du cycle 1.1 ont été recueillies du 1^{er} septembre 2000 au 3 novembre 2001 auprès d'un échantillon de 131 535 personnes. Le taux de réponse a été de 84,7 %.

Les participants à l'ESCC résidant au Québec ont été exclus de l'analyse car il est impossible de coupler les enregistrements de l'ISOP provenant du Québec à ceux de l'ESCC (22 667), parce que les dossiers d'hospitalisation fournis à Statistique Canada par cette province contiennent des numéros d'assurance-maladie brouillés, des codes postaux incomplets et ne donnent aucune date de naissance.

Toutes les données de l'ESCC, y compris les numéros d'assurance-maladie (NAM) provinciaux et les codes postaux, sont autodéclarées par les répondants, et l'ampleur de l'erreur pour ces variables est inconnue. Cependant, les applications de saisie de données qu'utilisent les intervieweurs comprennent des fonctions de dépistage des réponses incohérentes ou non incluses dans la fourchette de valeurs permises et des séquences alphanumériques non valables. D'autres renseignements sur l'ESCC figurent dans un autre rapport (Béland, 2002). On a demandé aux participants à l'ESCC s'ils autorisaient le couplage de l'information recueillie durant l'interview à celle figurant dans leur dossier de santé provincial, dont le recours passé et présent à des services comme ceux des hôpitaux, des cliniques ou des cabinets de médecin ou à d'autres services fournis par la province; neuf personnes sur dix ont accordé l'autorisation demandée. L'échantillon sur lequel porte la présente étude comprend 72 354 (66,5 %) personnes de 12 ans et plus réparties dans toutes les provinces et tous les territoires sauf le Québec, qui ont consenti au couplage et fourni un numéro d'assurance-maladie (NAM) valide (Tableau 2.1.1-1).

Tableau 2.1.1-1				
Nombre et pourcentage de participants à l'Enquête sur la santé dans les collectivités canadiennes qui ont consenti à ce que leurs réponses soient couplées aux données de leur dossier de santé administratif et qui ont fourni un NAM valide, selon certaines caractéristiques, Canada, Québec non compris, 2000-2001				
	Consentement au couplage		Consentement au couplage et NAM valide	
	Nombre	%	Nombre	%
Total	98 450	90,4	72 354	66,5
Province/territoires				
Terre-Neuve-et-Labrador	3 533	91,3	2 933	75,8
Île-du-Prince-Édouard	3 238	88,7	2 236	61,2
Nouvelle-Écosse	4 938	92,8	4 108	77,2
Nouveau-Brunswick	4 634	92,8	3 746	75,0
Ontario	35 674	90,8	24 917	63,4
Manitoba	7 653	90,4	5 552	65,5
Saskatchewan	7 417	92,6	6 142	76,7
Alberta	12 757	88,2	9 155	63,3
Colombie-Britannique	16 493	90,1	11 990	65,5
Territoires	2 113	83,9	1 575	62,6
Sexe				
Féminin	52 865	90,5	40 334	69,1
Masculin	45 585	90,3	32 020	63,4
Groupe d'âge				
12 à 74 ans	89 927	90,5	65 824	66,3
75 ans et plus	8 523	89,5	6 530	68,6

Source : Enquête sur la santé dans les collectivités canadiennes de 2000-2001.

Des poids de sondage ont été utilisés pour que les estimations calculées d'après les données de l'ESCC soient représentatives de la population cible plutôt que du seul échantillon. Le poids de sondage est le nombre de membres de la population étudiée que représente chaque répondant. Les poids de sondage reflètent les différentes probabilités de sélection et de réponse. Par conséquent, chaque enregistrement est pondéré par l'inverse de la probabilité de sélectionner la personne et d'obtenir sa réponse (Statistique Canada, 2008a). Le couplage d'enregistrements nécessite des poids de sondage supplémentaires, car les répondants ne consentent pas tous au couplage et ceux qui le font ne fournissent pas tous un NAM valide. Pour les besoins de la présente étude, nous avons calculé des poids de sondage corrigés pour tenir compte du consentement au couplage et de la déclaration d'un NAM valide.

Statistique Canada n'a pas accès aux bases de données provinciales sur l'assurance-maladie qui permettraient de vérifier les NAM déclarés par les participants à l'ESCC. À la place, chaque province et territoire fournit une formule de contrôle numérique qui est utilisée pour confirmer que les NAM sont au moins vraisemblables. Si les formules de contrôle ne peuvent pas remplacer les bases de données qui contiennent noms et prénoms, dates de naissance, adresses et NAM, elles permettent quand même de déceler les erreurs de transcription, comme l'inversion de deux chiffres, et constituent un moyen simple de faire la distinction entre des numéros valables et des séries de chiffres aléatoires.

2.1.2 Données hospitalières

La base de données de l'Information-santé orientée vers la personne (ISOP) tenue à jour par Statistique Canada contient des renseignements sur les départs de l'hôpital (congé et décès à l'hôpital) pour la quasi-totalité des hôpitaux de soins actifs et pour quelques hôpitaux psychiatriques, hôpitaux pour malades chroniques et établissements de réadaptation.

L'ISOP est un ensemble de données au niveau de la personne créé d'après les enregistrements des congés (parfois plusieurs congés chez une même personne) contenus dans la Base de données sur la morbidité hospitalière (BDMH). Les antécédents d'hospitalisation de chaque patient peuvent être établis d'après les enregistrements séquentiels au niveau de la personne de l'ISOP. Durant le processus de couplage, les enregistrements se rapportant à une même personne sont repérés au moyen du NAM, d'information démographique et de renseignements sur les diagnostics et les interventions (sexe, date de naissance, procédures propres à un sexe, et autres) (Division des méthodes d'enquêtes auprès des ménages, 2006).

Les dossiers d'hospitalisation de l'exercice écoulé sont ajoutés chaque année à l'ISOP. Chaque fois que les données d'une année sont entrées dans le système, le processus complet de l'ISOP est exécuté de nouveau afin de garantir la cohérence interne de l'information démographique au niveau de la personne sur les patients qui ont reçu de multiples congés.

Des études de seconde saisie des données, réalisées pour confirmer l'exactitude des dossiers d'hospitalisation, ont démontré la grande qualité des éléments non médicaux des données administratives, essentielles au couplage d'enregistrements. Ainsi, 99 % d'un échantillon aléatoire des enregistrements de congé relatifs aux séjours à l'hôpital survenus de septembre à la fin de novembre 2000 contenaient un NAM exact et 91 %, un code postal sans erreur (Richards et coll., 2001).

Statistique Canada possède des données hospitalières comprenant les NAM pour toutes les provinces (sauf le Québec) et les Territoires du Nord-Ouest de 1994-1995 à ce jour, ainsi que pour certaines provinces pour 1992-1993 et 1993-1994. Même si la base de données de l'ISOP contient la grande majorité des enregistrements de la BDMH, environ 3 % correspondant à des personnes de 12 ans et plus (la population cible de la présente étude) en sont exclus parce que le NAM manque ou n'est pas valide (Division des méthodes d'enquêtes auprès des ménages, 2006).

Du 1^{er} septembre 2000 au 3 novembre 2001, les hôpitaux de soins actifs hors Québec ont accordé 2,3 millions de congés à 1 624 972 personnes de 12 ans et plus. L'étude ne porte pas sur les congés d'autres hôpitaux que ceux de soins actifs, car leur couverture n'est pas uniforme d'une province à l'autre.

Les populations cibles de l'ESCC et de l'ISOP diffèrent quelque peu. L'ESCC exclut les membres à plein temps des Forces canadiennes de même que les personnes vivant dans les réserves indiennes, en établissement (p. ex. foyers de soins infirmiers et prisons) et dans certaines régions éloignées. Par contre, l'ISOP est un recensement, de sorte que les hospitalisations englobent ces populations. Afin d'établir une meilleure concordance entre les populations de l'ESCC et de l'ISOP, les hospitalisations qu'il a été possible d'attribuer à des personnes vivant dans les réserves ou en établissement ont été exclues de l'analyse.

La population des réserves est une variable de recensement dérivée, obtenue en déterminant le type de subdivision de recensement (SDR) conformément aux critères établis par Affaires indiennes et du Nord Canada (AINC). Elle comprend aussi la population de certaines SDR correspondant aux collectivités du Nord de la Saskatchewan, des Territoires du Nord-Ouest et du Yukon (Statistique Canada, 2002). Les personnes vivant en établissement ont été identifiées à l'aide du Fichier de conversion des codes postaux (FCCP+) (Statistique Canada, 2008c) et d'une liste d'établissements servant à l'Enquête sur les établissements de soins pour bénéficiaires internes (Statistique Canada, 2008b). Les hospitalisations de 31 330 personnes vivant dans les réserves et sur les terres connexes ont été éliminées de l'ISOP, ainsi que celles de 21 299 personnes vivant en établissement. La suppression de ces 52 629 enregistrements, soit environ 3 % des personnes hospitalisées selon l'ISOP au cours de la période étudiée, a fait mieux concorder la population couverte par l'ISOP à la population cible de l'ESCC.

2.2 Techniques d'analyse

2.2.1 Couplage d'enregistrements probabiliste

Les participants à l'ESCC qui avaient été hospitalisés ont été identifiés par couplage d'enregistrements probabiliste. Le couplage des données de l'ESCC et de l'ISOP a été exécuté au moyen du Système généralisé de couplage d'enregistrements (SGCE), un logiciel développé par Statistique Canada. Les deux sources de données contiennent de nombreuses variables, mais n'ont en commun que quelques champs qui sont suffisamment distincts pour être

utilisés pour effectuer l'appariement. Nous avons considéré qu'un participant à l'ESCC avait été hospitalisé si l'ISOP contenait un enregistrement comprenant un NAM et (ou) des caractéristiques démographiques comparables (date de naissance, sexe, code postal), ainsi qu'une date d'admission dans un établissement de soins actifs entre le 1^{er} septembre 2000 et le 3 novembre 2001.

Le couplage probabiliste ne nécessite pas la concordance complète des variables d'appariement. La qualité de l'appariement d'une paire d'enregistrements est plutôt évaluée par l'application d'algorithmes qui mesurent la probabilité d'un appariement correct (Fair et Whitridge, 1997, Fellegi et Sunter, 1969). Des points ont été attribués ou déduits, selon la similitude des valeurs des champs. Par exemple, un score positif élevé a été attribué si les NAM étaient identiques et que la province de délivrance était la même; si les valeurs étaient comparables, sans toutefois être exactes, un score positif plus faible a été accordé pour tenir compte de la concordance partielle; si les valeurs contenues dans les deux enregistrements étaient entièrement différentes, des points ont été soustraits.

Le nombre de points attribués à chaque paire de variables de couplage reflétait leur importance en tant que variables d'appariement, laquelle était normalement fonction de leur unicité. Ainsi, comme il n'existe que deux valeurs possibles du sexe d'un répondant ou d'un patient, le nombre de points attribués à la concordance des valeurs de cette variable était inférieur à celui octroyé si les codes postaux ou les NAM concordaient.

Le score de couplage total attribué à chaque paire d'enregistrements ESCC-ISOP a été calculé en additionnant les points alloués à chaque paire de variables d'appariement. Il était d'autant plus probable que les deux enregistrements concernent la même personne que le score de couplage total était élevé. Les poids de couplage totaux suivent idéalement une loi bimodale. Les paires d'enregistrements dont le score était supérieur au seuil choisi ont été considérées comme des appariements vrais, tandis que celles dont le score était inférieur au seuil établi ont été rejetées. Pour ne pas avoir à procéder à un examen manuel, les seuils d'exclusion choisis pour l'étude étaient identiques, si bien qu'un seul résultat (appariement et non-appariement) pouvait être attribué à chaque paire d'enregistrements.

3. Résultats

Afin d'évaluer la couverture du couplage des données du cycle 1.1 de l'ESCC à celles de l'ISOP, nous avons comparé le nombre de personnes admises à l'hôpital selon chaque source de données. Nous avons appliqué des poids de sondage, corrigés pour tenir compte du consentement au couplage et de la validité du NAM, aux enregistrements des participants à l'ESCC pour lesquels un enregistrement figurait aussi dans la base de données de l'ISOP. Nous avons considéré le nombre d'hospitalisations selon l'ISOP comme étant la norme. Pour calculer le taux de couverture, nous avons divisé les estimations pondérées des participants à l'ESCC pour lesquels un couplage avait eu lieu avec un enregistrement de l'ISOP par le nombre selon l'ISOP, dont étaient soustraits les enregistrements des personnes vivant dans les réserves indiennes et sur les terres connexes ou en établissement, puis nous avons multiplié le quotient par 100. Nous avons ensuite examiné les écarts entre les chiffres selon l'ISOP et les estimations pondérées fondées sur l'ESCC. Les erreurs-types et les intervalles de confiance à 95 % des taux de couverture ont été calculés par la méthode du bootstrap et la signification statistique a été déterminée au moyen d'un test t ($p < 0,05$) (Rao et coll., 1992, Rust et Rao, 1996).

Selon l'ISOP, du 1^{er} septembre 2000 au 3 novembre 2001, 1 572 343 personnes ont été admises à un hôpital de soins actifs (excluant ceux du Québec) (tableau 3-1). Les estimations pondérées fondées sur l'ESCC, corrigées pour tenir compte du consentement au couplage et d'un NAM valide, sont inférieures de 7,7 % à ce chiffre (1 451 272).

Les taux de couverture sont comparables pour les hommes et pour les femmes (91,0 % et 93,1 %). Les taux provinciaux ne diffèrent pas significativement du taux pour le reste du Canada. Toutefois, le nombre d'habitants des territoires ayant été hospitalisés estimé selon l'ESCC est considérablement plus élevé que celui donné par l'ISOP. Par conséquent, le taux de couverture des territoires est supérieur à 100 %.

Les taux de couverture sont comparables pour la plupart des groupes d'âge. Font exception les personnes de 75 ans et plus dont le taux (76,2 %) est significativement plus faible que celui calculé pour le groupe des 12 à 74 ans (96,4 %).

Tableau 3-1

Nombre de personnes admises dans les hôpitaux de soins actifs et taux de couverture, du 1^{er} septembre 2000 au 3 novembre 2001, selon certaines caractéristiques et la source des données, population de 12 ans et plus, Canada, Québec non compris

	Information-santé orientée vers la personne (ISOP) Nombre	Enquête sur la santé dans les collectivités canadiennes (ESCC)		Taux de couverture ESCC/ISOP Intervalle de confiance à 95 %		
		Nombre non pondéré	Nombre pondéré	Taux %	de	à
Total	1 572 343	6 785	1 451 272	92,3	88,9	95,7
Province/territoires						
T.-N.-L.	41 394	272	40 445	97,7	83,6	111,8
Î.-P.-É.	11 784	237	11 061	93,9	79,6	108,1
N.-É.	67 226	348	60 419	89,9	78,0	101,7
N.-B.	67 542	423	62 203	92,1	81,7	102,5
Ont.	753 970	2 230	694 463	92,1	86,6	97,6
Man.	82 386	567	69 739	84,6	73,6	95,7
Sask.	82 778	659	78 664	95,0	86,4	103,7
Alb.	202 498	863	186 301	92,0	83,3	100,7
C.-B.	258 883	1 062	241 647	93,3	85,3	101,3
Territoires	3 882	124	6 331	163,1*	139,3	186,9
Sexe						
Féminin [†]	971 087	4 343	904 318	93,1	88,8	97,5
Masculin	601 249	2 442	546 955	91,0	85,4	96,5
Groupe d'âge						
12 à 74 ans [†]	1 252 336	5 299	1 207 392	96,4	92,4	100,4
75 ans et plus	320 007	1 486	243 881	76,2*	70,2	82,2

[†] catégorie de référence
* valeur significativement différente de celle observée pour la catégorie de référence (p<0,05); pour les comparaisons provinciales, valeur significativement différente de celle observée pour le reste du Canada, par exemple, Ontario comparé au Canada moins l'Ontario.
Source : Enquête sur la santé dans les collectivités canadiennes, 2001 et Information-santé orientée vers la personne, 2000-2001 à 2001-2002.

4. Discussion

Le taux de couverture significativement plus faible des personnes de 75 ans et plus était attendu, car les deux sources de données ne se rapportent pas exactement aux mêmes populations. L'ESCC exclut les personnes vivant en établissement, lesquelles sont incluses dans les données hospitalières (ISOP). Or, le placement en établissement est beaucoup plus fréquent chez les personnes âgées que chez les jeunes : dans l'ensemble, moins de 2 % des Canadiens vivent en établissement, mais chez les 75 ans et plus, la proportion passe à 16 % (Turcotte et Schellenberg, 2007).

À défaut de renseignements directs dans les enregistrements de l'ISOP sur le lieu de résidence des patients, le code postal conjugué au FCCP+ et à la liste des établissements de soins pour bénéficiaires internes a servi à déterminer si ceux-ci vivaient en établissement. Plus de 20 000 résidents d'établissements ont ainsi été identifiés et supprimés de l'ISOP. Néanmoins, le taux de couverture des personnes de 75 ans et plus est demeuré significativement inférieur aux taux calculés pour les personnes plus jeunes.

L'utilisation du FCCP+ et de la liste des établissements de soins pour bénéficiaires internes en vue de repérer les établissements uniquement d'après le code postal n'est pas la méthode idéale. Ceux où vivait la majorité de la population ayant un même code postal étaient plus susceptibles que les autres d'être repérés, puis éliminés des données de l'ISOP. En outre, les codes postaux des établissements situés en milieu urbain sont plus précis de sorte que leurs résidents étaient plus susceptibles d'être éliminés de l'ISOP. Souvent, dans les régions rurales, les banlieues éloignées et les petites villes, le même code postal est attribué à plusieurs secteurs de dénombrement et aires de diffusion. Leur codage est donc beaucoup moins précis que les codes postaux urbains centralisés, qui sont normalement rattachés à un seul secteur de dénombrement ou à une seule aire de diffusion. Il s'ensuit que les dénombrements fondés sur l'ISOP ont probablement continué d'inclure les résidents des établissements des régions rurales, des banlieues éloignées et des petites villes.

Dans les territoires, le taux de couverture pose aussi un problème, étant donné que les estimations fondées sur les données couplées ESCC-ISOP étaient supérieures à la norme (ISOP). Toutefois, cette situation est moins préoccupante, car le faible nombre d'enregistrements de l'ESCC qui ont été couplés à l'ISOP (124) écarte la possibilité que de futures analyses portent sur cette sous-population. Le taux de couverture des territoires était de 113 % avant que la population des réserves soit éliminée des données de l'ISOP, après quoi il est passé à 163 %. Les raisons pour lesquelles l'estimation calculée d'après les données couplées de l'ESCC est beaucoup plus élevée que celle basée sur l'ISOP ne sont pas claires. Les enregistrements des participants à l'ESCC habitant les territoires ont été examinés afin de voir si un poids de sondage élevé avait été attribué à certains d'entre eux, ce qui pourrait expliquer l'écart entre les dénombrements selon l'ISOP et ceux issus du couplage ESCC-ISOP. Aucun poids divergent n'a été découvert : la moyenne était de 51, et les poids se situaient tous dans la fourchette de 11 à 178. Par ailleurs, il a été impossible de relever les hospitalisations de membres du personnel militaire afin de les éliminer de l'ISOP. Les membres à temps plein des Forces canadiennes étant exclus de l'ESCC, leur prise en compte dans l'ISOP pourrait donc influencer sur le taux de couverture.

5. Conclusion

La valeur du couplage d'enregistrements pour les études épidémiologiques de la santé des populations n'est plus à démontrer. L'appariement d'information tirée de données administratives sur la santé recueillies régulièrement, comme celles de l'ISOP, et de données d'enquête, comme celles de l'ESCC, promet des découvertes au sujet des déterminants de la santé, des diverses catégories d'utilisation des soins de santé et des résultats en matière de santé. L'évaluation de la couverture est une condition préalable fondamentale aux analyses qui regroupent des données sur la santé provenant de sources multiples en s'appuyant sur le fichier couplé ESCC-ISOP.

La présente évaluation montre que le taux de couverture global est élevé, souvent supérieur à 90 %, bien qu'il soit plus faible chez certains participants à l'ESCC, notamment les personnes âgées. Cependant, cette limite n'est pas un obstacle, pourvu que les utilisateurs du fichier reconnaissent explicitement que les conclusions ne concernent que la population à domicile générale (population cible de l'ESCC) et non la population totale, particulièrement les personnes vivant en établissement.

Remerciements :

L'auteure remercie Claude Nadeau de son aide et de ses judicieux commentaires.

Bibliographie

Béland, Y. (2002). Enquête sur la santé dans les collectivités canadiennes - aperçu de la méthodologie, *Rapports sur la santé* 13, 9-15.

Division des méthodes d'enquêtes auprès des ménages. (2006). External linkage production report: Data years: F1992 to F2004, rapport inédit. Statistique Canada: Ottawa, Canada.

- Fair, M. (2004). Generalized Record Linkage System - Statistics Canada's Record Linkage Software, *Austrian Journal of Statistics*; 33, 37-53.
- Fair, M.E. et Whitridge, P. (1997). Tutorial on Record Linkage, *Federal Committee on Statistical Methodology*. www.fcsm.gov/working-papers/RL_chap12.html.
- Fellegi, I.P. et Sunter, A.B. (1969). A Theory for Record Linkage, *Journal of the American Statistical Association*, 64, 1183-210.
- Nadeau, C., Beaudet, M.P. et Marion, J. (2006). Couplage d'enregistrements déterministe et probabiliste, *Recueil: Symposium 2006, Enjeux méthodologiques reliés à la mesure de la santé des populations*. Statistique Canada: Ottawa, Canada.
- Rao, J.N.K., Wu, C.F.J., et Yue, K. (1992), Quelques travaux récents sur les méthodes de rééchantillonnage applicables aux enquêtes, *Techniques d'enquête*, 18, 225-234.
- Richards, J., Brown, A., et Homan, C. (2001), L'étude de la qualité des données de la base canadienne de données sur les congés des patients, *Recueil: Symposium 2001, La qualité des données d'un organisme statistique : une perspective méthodologique*. Statistique Canada: Ottawa, Canada. www.statcan.ca/english/freepub/11-522-XIE/2001001/session16/s16a.pdf.
- Rust, K.F. et Rao, J.N.K. (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Statistical Methods in Medical Research*, 5, 281-310.
- Statistique Canada, (2001). Dictionnaire du Recensement de 2001, Statistique Canada: Ottawa, Canada. <http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=fra&catno=92-378-X>
- Statistique Canada (2008a). Enquête sur la santé dans les collectivités canadiennes — Estimation, Statistique Canada: Ottawa, Canada. www.statcan.gc.ca/cgin/imdb/p2SV.pl?Function=getSurvey&SurvId=3226&SurvVer=0&InstaId=15282&InstaVer=1&SDDS=3226&lang=en&db=imdb&adm=8&dis=2#b7
- Statistique Canada (2008b). Établissements de soins pour bénéficiaires internes, Statistique Canada: Ottawa, Canada. <http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=83-237-X&lang=fra>
- Statistique Canada (2008c). Fichier de conversion des codes postaux (FCCP): mise à jour, Statistique Canada: Ottawa, Canada. <http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=92-153-UCB&lang=fra>.
- Turcotte M. et Schellenberg G. (2007). *Un portrait des aînés au Canada*, Statistique Canada: Ottawa, Canada.