Les indicateurs socioéconomiques régionaux : des outils de recherche, de politiques et de planification axés sur les disparités d'état sanitaire

Keith Denny, Ph.D.¹, Mélanie Josée Davidson, B.A.²

Ce numéro spécial de la *Revue canadienne de santé publique* porte sur l'utilisation contemporaine des indicateurs socioéconomiques régionaux, ou ISR. Comme leur nom l'indique, les ISR sont des outils d'analyse d'information (provenant souvent du recensement) spécifiquement sélectionnés pour caractériser le profil socioéconomique de zones géographiques plutôt que de personnes. Les indicateurs à une seule composante, comme le revenu selon le quartier, en sont des exemples courants; les indicateurs de défavorisation régionaux à plusieurs composantes sont un autre exemple d'ISR.

L'intérêt pour de tels instruments n'est pas aussi récent qu'on pourrait le croire. Dès la seconde moitié du XIX^e siècle, on emploie en Angleterre les outils de ce qui constitue à l'époque la nouvelle science de la statistique pour décrire la *géographie* de la santé, notamment en illustrant les écarts de mortalité entre les zones urbaines et rurales, mais aussi entre les agglomérations urbaines riches et pauvres^{1,2}. Au début du XX^e siècle, aux États-Unis, où les orientations empiriques fondées sur la statistique commencent à s'imposer dans les sciences sociales, on fait de multiples tentatives pour catégoriser objectivement ce qu'on appelle déjà le statut socioéconomique (SSE)³⁻⁶. Chapin, par exemple, met au point une échelle à quatre composantes (les possessions culturelles, le revenu effectif, la participation à des activités de groupe et les possessions matérielles) « pour permettre l'équation de ces facteurs dans des études expérimentales » (traduction libre, note 4, p. 99). C'est un indice qui a quelque chose de familier près d'un siècle plus tard.

Dans les années 1930, toujours aux États-Unis, cet intérêt pratique pour la statistique et la géographie mène à la première utilisation des secteurs de recensement dans les analyses sociales^{7,8}; et à la fin des années 1950, Coulter et Guralnick rapportent que les données du recensement sont utilisées au niveau des secteurs de recensement, conjointement avec les statistiques démographiques, pour décrire les écarts dans les taux de mortalité de groupes socioéconomiques⁹. Leurs contemporains Farber et Osoinach (1959) mettent au point un indice du rang socioéconomique des secteurs de recensement à l'aide des données du recensement sur la profession, l'instruction, le revenu et la population non blanche¹⁰.

Une grande partie de cette histoire reste inconnue aujourd'hui, mais il ne fait aucun doute que les ISR suscitent un intérêt croissant depuis une vingtaine d'années. Cet intérêt a trois raisons distinctes, mais intimement liées : les lacunes dans les données (de micro-niveau), l'émergence d'une préoccupation conceptuelle et théorique pour le « lieu » et la santé, et le nouvel impératif d'ancrer la recherche dans les inégalités sociales de santé afin de faciliter la production de preuves pouvant étayer des interventions fondées sur le lieu.

Les chercheurs et les planificateurs canadiens qui s'emploient à décrire et à corriger les disparités socioéconomiques dans la santé sont sans cesse confrontés à un épineux problème : le manque de données disponibles englobant à la fois la santé et l'information socioéconomique. Peu de statistiques démographiques canadiennes incluent des données socioéconomiques qui faciliteraient la localisation des disparités sur le plan des issues de la grossesse et de la mortalité aux États-Unis, par exemple, les registres des naissances contiennent des données sur le niveau d'instruction maternel, mais seules deux pro-

vinces canadiennes recueillent ce type d'information : Québec et Terre-Neuve-et-Labrador¹¹. Les registres des décès aux États-Unis précisent la profession et le niveau d'instruction¹². Les données sur la profession servent principalement à la recherche en santé portant sur ce sujet (voir par exemple la note 13). Au Royaume-Uni, la profession est aussi enregistrée sur le certificat de décès; à la différence des États-Unis, les catégories professionnelles y servent depuis près d'un siècle à établir le « schéma des classes sociales » du registraire général de l'état civil. Ce schéma permet des descriptions détaillées des écarts socioéconomiques de mortalité¹⁴⁻¹⁶. Le schéma du registraire général de l'état civil et les analyses qui s'en inspirent ont leur part de critiques¹⁷, mais au Canada, on ne peut même pas envisager de telles recherches, car bien que les dossiers de mortalité incluent de l'information sur le sexe, l'âge et la cause du décès, ils ne contiennent aucune donnée pouvant servir à déterminer le statut socioéconomique.

Comme pour les statistiques démographiques, la plupart des données – tant administratives que cliniques – générées par le système de santé et les registres des maladies souffrent de la même absence de contenu socioéconomique. L'information du secteur de la santé, comme les registres des hospitalisations, inclut rarement des données sur le revenu, l'instruction, la profession ou d'autres indicateurs socioéconomiques. Certains fournisseurs de soins de santé primaires et de soins à domicile ou prolongés peuvent recueillir des données démographiques et socioéconomiques sur leurs patients et usagers, mais il reste qu'il n'y a pas de système, au Canada, pour recueillir, normaliser et agréger systématiquement ce genre d'information.

Les enquêtes sur la santé de la population peuvent se prêter à des analyses de haut niveau sur les inégalités sociales de santé, dont la morbidité et l'invalidité, et sur l'accès et l'utilisation des soins de santé. L'Enquête sur la santé dans les collectivités canadiennes (ESCC), par exemple, rassemble des données sanitaires et socioéconomiques; c'est l'une des principales sources d'information pour la santé publique au Canada. Toutefois, malgré la taille importante de son échantillon, l'ESCC est conçue pour être utilisée au niveau des régions sanitaires et non pour l'analyse sous-régionale de petits territoires. À l'échelle des petites zones géographiques, les échantillons ne sont souvent pas assez gros pour produire des renseignements fiables, même lorsqu'il est possible de fusionner plusieurs cycles de l'Enquête. L'ESCC permet de mieux comprendre les liens entre le SSE individuel et la santé, mais elle n'est pas assez locale pour donner de l'information sur les contextes où ces liens se manifestent.

De telles contraintes constituent des lacunes importantes pour la détection, la compréhension et la surveillance des disparités d'état sanitaire.

Les ISR – dérivés des zones géographiques – peuvent être utilisés conjointement avec les statistiques démographiques et les données administratives du système de santé pour permettre un compromis méthodologique qui remplace les données de microniveau (par per-

Affiliations des auteurs

1. Directeur de la recherche et des politiques, Association canadienne des soins de santé, Ottawa (Ontario); professeur-chercheur adjoint, Département de sociologie et d'anthropologie, Université Carleton, Ottawa (Ontario)
2. Institut canadien d'information sur la santé, Ottawa (Ontario)

sonne ou par ménage) par des données agrégées. Ce genre d'analyse implique d'habitude l'appariement de particuliers à un emplacement spatial à l'aide des renseignements sur le lieu de résidence, comme les codes postaux, qui sont habituellement recueillis dans les dossiers administratifs et cliniques*. Le recours à ces renseignements permet de compiler des données sanitaires, sociodémographiques et autres pour créer des profils par petite zone (par exemple, une aire de diffusion ou un secteur de recensement).

Utilisés de cette façon, les ISR servent de compromis opérationnel aux caractéristiques individuelles dont ils sont les substituts. Pour que la substitution soit valide, il faut être raisonnablement sûr que les zones mêmes sont suffisamment homogènes sur le plan socioéconomique (et définies à une échelle qui permet une telle homogénéité), afin que les données agrégées dépeignent une personne composite ou moyenne²⁰. Autrement dit, plus une zone est homogène, plus la substitution des données de microniveau par les données de la zone est fiable. Ce besoin d'être assuré d'un haut degré d'homogénéité signifie que l'on considère généralement l'utilisation des ISR comme convenant mieux à l'analyse d'agglomérations urbaines, où la densité de population et l'homogénéité ont tendance à être plus élevées.

L'une des difficultés constantes de l'utilisation des ISR a trait précisément à la mesure dans laquelle (nonobstant l'homogénéité estimée de la zone) les variables agrégées qui les composent *sont bel et bien* des substituts raisonnables aux caractéristiques individuelles²¹. Certains chercheurs, comme Geronimus, constatent « des effets estimatifs considérablement plus importants, quand on utilise des variables agrégées, que lorsqu'on utilise des variables de microniveau » (traduction libre, note 21, p. 835), tandis que d'autres concluent que « le risque, s'il en est, en l'absence d'information socioéconomique individuelle, est d'obtenir une estimation prudente des inégalités socioéconomiques de santé » (traduction libre, note 22, p. 833; voir aussi les notes 23 et 24).

Une autre tendance qui favorise l'utilisation des ISR est l'intérêt conceptuel croissant pour les influences du contexte (ou du lieu) sur la santé²¹. Dans ce cas, ce n'est pas simplement par manque de données individuelles que l'on utilise des variables régionales pour l'analyse. Selon Subramanian et coll., par exemple, les ISR saisissent en fait des phénomènes que les données de microniveau ne saisissent pas : « une dimension collective ou contextuelle commune *que partagent toutes les personnes dans un certain contexte* » (traduction libre, note 22, p. 833, nos italiques).

Certains chercheurs préconisent une perspective méthodologique plus nuancée, qui n'insiste pas pour établir une distinction claire entre les phénomènes régionaux et de microniveau, car l'exercice peut être futile. Comme l'écrit Geronimus, « il est difficile, sur le plan conceptuel et empirique, de distinguer les impacts des caractéristiques individuelles et communautaires en raison de la complexité des processus sociaux en cause, et pour la raison évidente qu'on n'affecte pas aléatoirement des particuliers à un lieu de résidence ou à une identité sociale » (traduction libre, note 21, p. 839). Nous aurions tort, pense-t-il, de prendre pour acquis que ce que nous mesurons à l'aide de données de microniveau, ce sont *exclusivement* des caractéristiques individuelles, et que les variables agrégées ne reflètent *que* des phénomènes contextuels.

Macintyre et coll. vont plus loin en faisant observer que « la distinction entre la composition [les variables *individuelles* agrégées] et le contexte [les variables *régionales* supra-individuelles] pourrait ne

pas être aussi claire, sur le plan conceptuel, ni aussi utile qu'il peut sembler à première vue. Les propriétés des particuliers ou des ménages utilisées dans de nombreux modèles multiniveaux sont elles-mêmes façonnées par les propriétés de la localité utilisée dans les mêmes modèles » (traduction libre, note 25, p. 129). Ces chercheurs font remarquer que l'emploi et la profession, en tant que variables de microniveau, sont façonnés et limités dans les situations sociales concrètes par l'économie du marché du travail local, tout comme le niveau d'instruction individuel s'inscrit dans le contexte accidentel des normes, des milieux et des ressources des écoles locales²⁵.

S'il y a des raisons convaincantes d'utiliser des variables régionales – autrement que pour remplacer des données de microniveau manquantes – on peut se demander s'il vaut vraiment mieux saisir les phénomènes régionaux les plus importants (du point de vue de la santé) au moyen des variables couramment utilisées dans les ISR alors que ces variables sont recueillies pour fournir de l'information sur les particuliers et les ménages, et sélectionnées selon leur disponibilité, à partir d'hypothèses sur leur pouvoir de substitution. Une autre question a trait au niveau d'agrégation proprement dit.

Il y a deux aspects à la question de l'échelle géographique. Le premier est un aspect méthodologique, brièvement abordé plus tôt, que l'on appelle en anglais le « problème de l'unité aérienne modifiable ». Il s'agit du biais qui peut s'introduire dans l'analyse lorsqu'on emploie des unités d'analyse qui résultent d'impératifs administratifs ou politiques plutôt que de principes empiriques rationnels²⁰. Le deuxième aspect consiste à savoir si les unités géographiques analysées ont un lien quelconque avec des communautés réelles (ici encore, en raison de leurs origines administratives). Kreiger et coll. font remarquer que « la validité de l'utilisation d'indicateurs socioéconomiques régionaux dépend de la mesure dans laquelle les régions en question constituent des unités géographiques rationnelles – ce qui est différent de savoir si elles sont des “substituts” aux données socioéconomiques individuelles » (traduction libre, note 23, p. 479). Il est sous-entendu que les frontières géographiques politico-administratives du recensement ou des codes postaux, utilitaires de nature, pourraient ne pas être le moyen le plus rationnel de diviser le monde afin de mieux comprendre les déterminants de la santé, ni inciter à s'attaquer aux iniquités de santé. Macintyre et coll. expriment la même préoccupation de la façon suivante : « le simple fait qu'un territoire est d'une taille appropriée pour un enquêteur du recensement ou pour la personne qui livre le courrier ne veut pas dire qu'il est à une échelle appropriée pour l'analyse [des effets socioéconomiques sur la santé] » (traduction libre, note 25, p. 134). Autrement dit, même si nous concédons que les ISR ont une valeur, au-delà de leur utilisation comme solution de rechange en l'absence de données de microniveau, il reste, comme Foley l'a fait remarquer il y a plus d'un demi-siècle, que « la plus grande valeur des données des secteurs de recensement, pour la recherche en général, est peut-être qu'elles fournissent des *profils écologiques bruts* » (traduction libre, note 8, p. 738, nos italiques).

Il y a une troisième raison derrière l'utilisation croissante des ISR : bien qu'ils représentent un compromis opérationnel et méthodologique pour la mesure des disparités sociales et de la santé, leur utilisation présente plusieurs avantages pratiques. En faisant preuve de prudence, planificateurs et gestionnaires peuvent tirer parti d'analyses appropriées et adaptées aux besoins de la prise de décisions locales, analyses auxquelles ils n'auraient pas accès autrement. De plus, les ISR peuvent être utilisés conjointement avec les systèmes d'information géospatiale (SIG) et les logiciels de cartographie pour créer des images parlantes de la distribution sociale de la santé et des facteurs

* Les travaux de Wilkins et coll. ont beaucoup aidé les chercheurs du Canada à relier facilement les codes postaux aux unités géographiques normalisées du Recensement (voir par exemple les notes 18 et 19).

liés à la santé. De tels outils peuvent à tout le moins fournir des indications utiles sur les endroits à étudier de plus près; elles peuvent aussi favoriser les débats entre les acteurs locaux en faisant référence à des lieux concrets plutôt qu'à des groupements de population abstraits.

En plus de l'indicateur de revenu selon le quartier de Statistique Canada, largement utilisé et à une seule composante, on observe depuis les 20 dernières années un intérêt croissant pour l'élaboration et l'utilisation d'ISR à plusieurs composantes dans les études sur la santé des populations. Ce numéro spécial de la *Revue canadienne de santé publique* réunit pour la première fois des descriptions d'expériences canadiennes menées avec des ISR à plusieurs composantes et donne un éclairage sur leur élaboration et leur utilisation. Les articles dans ce numéro s'adressent directement à un lectorat de planificateurs et de décideurs canadiens de première ligne; c'est pourquoi on y trouvera surtout des aperçus descriptifs de l'élaboration et de l'application d'ISR, avec leurs forces et leurs limites, ainsi qu'une discussion de la mesure dans laquelle ils sont utilisés pour formuler des politiques et planifier des programmes.

Pour mettre les choses en contexte, ce numéro spécial s'ouvre sur l'exemple de la Nouvelle-Zélande, où l'indice de défavorisation néo-zélandais (NZDep) sert d'instrument de la politique nationale depuis 15 ans. Dans leur article, Salmond et Crampton décrivent comment on a élaboré le NZDep et comment on l'utilise pour visualiser les écarts de santé et affecter des fonds aux zones locales en fonction des besoins. Les auteurs soulignent que l'indice a évolué au fil du temps, d'après la rétroaction de ses utilisateurs, et nous éclairent sur certaines des difficultés de l'utilisation des ISR pour la recherche, la planification et les politiques.

Les quatre articles suivants portent sur des exemples canadiens d'ISR à plusieurs composantes. Ces ISR ont été élaborés en Ontario (CAN-Marg), au Québec (indice de défavorisation), au Manitoba (SEFI) et en Colombie-Britannique (VANDIX). Matheson et coll. présentent l'indice de marginalisation canadien (CAN-Marg), structuré selon quatre dimensions distinctes de la marginalisation. Ils décrivent l'élaboration de cet ISR et les associations de ses diverses composantes avec les disparités d'état sanitaire au Canada. Pampalon et coll. décrivent aussi un ISR à plusieurs composantes qui peut être utilisé dans tout le pays et montrent comment les deux dimensions (matérielle et sociale) de l'indice servent à la recherche et à la planification au Québec, en particulier. Bien que le CAN-Marg et l'indice de défavorisation mesurent différentes conceptualisations du statut socioéconomique, les deux ont été créés à l'aide d'analyses factorielles des données du Recensement et ont montré qu'ils peuvent révéler les inégalités sociales de santé.

Ensuite, Chateau et coll. présentent l'indice des facteurs socioéconomiques (SEFI et SEFI-2) et ses associations avec les résultats de santé de la population manitobaine. Créé à l'origine pour contourner le problème de la suppression des données sur le revenu à petite échelle, la nouvelle version de cet ISR (le SEFI-2) inclut des variables du Recensement liées à la défavorisation matérielle et sociale, dont le revenu, un peu comme l'indice de défavorisation employé au Québec. Il fait cependant appel à des variables de rechange pour mieux refléter les diverses dimensions de la défavorisation dans le contexte manitobain.

Enfin, Bell et Hayes présentent le VANDIX, un indice de défavorisation à l'échelle des quartiers de la région de Vancouver. Contrairement aux trois exemples précédents d'ISR présentés dans ce numéro, le VANDIX a été créé pour contourner le problème du manque de données sanitaires à petite échelle géographique à l'intérieur d'une ville (plutôt que d'essayer de pallier le manque d'information socioéconomique

dans les dossiers médicaux). Cet ISR sert donc d'indicateur indirect de la santé. Comme les autres ISR dans ce numéro, le VANDIX est fondé sur les données du Recensement, mais il a été élaboré à partir de sondages auprès d'informateurs au lieu d'être guidé par les données.

À la différence d'autres pays (le Royaume-Uni ou la Nouvelle-Zélande, par exemple), aucun indicateur régional du statut socioéconomique ou de la défavorisation ne s'est normalisé au Canada. Au lieu de cela, plusieurs outils provinciaux ont émergé, ce qui peut être le signe que c'est à ce niveau que la normalisation s'établira. Ce que tous les ISR canadiens ont en commun, toutefois, c'est qu'ils ont jusqu'ici puisé dans les données obtenues grâce au questionnaire détaillé (à participation obligatoire) pour le Recensement. Tous les effets du remplacement récent du questionnaire détaillé par l'Enquête nationale auprès des ménages (à participation volontaire) restent à déterminer. Entre-temps, les ISR contribuent à combler un écart important dans notre compréhension et notre surveillance des inégalités sociales de santé.

BIBLIOGRAPHIE

1. Rosen G. *A History of Public Health*. New York, Johns Hopkins University Press, 1958.
2. Porter D. *Health, Civilization and the State: A History of Public Health from Ancient to Modern Times*. Londres, Routledge, 1999.
3. Sydenstricker E, King W. The measurement of the relative income of families. *Quarterly Publications of the American Statistical Association* 1921;17(135):842-57.
4. Chapin FS. A quantitative scale for rating the home and social environment of middle class families in an urban community: A first approximation to the measurement of socio-economic status. *J Educ Psychol* 1928;19(2):99-111.
5. Lunberg GA. The measurement of socioeconomic status. *Am Sociol Rev* 1940;5(1):29-39.
6. Kahl JA, Davis JA. A comparison of indexes of socio-economic status. *Am Sociol Rev* 1955;20(3):317-25.
7. Green WG. The use of census tracts in analyzing the population of a metropolitan community. *J Am Stat Assoc* 1933;28(181):147-53.
8. Foley DL. Census tracts and urban research. *J Am Stat Assoc* 1953;48(264):733-42.
9. Coulter EJ, Guralnick L. Analysis of vital statistics by census tract. *J Am Stat Assoc* 1959;54(288):730-40.
10. Farber B, Osoinach JC. An index of socio-economic rank of census tracts in urban areas. *Am Sociol Rev* 1959;24(5):630-40.
11. Luo ZC, Wilkins R, Kramer MS. Effect of neighbourhood income and maternal education on birth outcomes: A population-based study. *JAMC* 2006;174(10):1415-21.
12. Bidulescu A, Rose KM, Wolf SH, Rosamond WD. Occupation recorded on certificates of death compared with self-report: The Atherosclerosis Risk in Communities (ARIC) Study. *BMC Public Health* 2007;7(229).
13. Dubrow R, Sestito JP, Lalich NR, Burnett CA, Salg JA. Death certificate-based occupational mortality surveillance in the United States. *Am J Ind Med* 1987;11(3):329-42.
14. Stevenson THC. The social distribution of mortality from different causes in England and Wales, 1910-12. *Biometrika* 1923;15(3/4):382-400.
15. Szreter S. The genesis of the registrar-general's social classification of occupations. *Br J Sociol* 1984;35(4):522-46.
16. Pamuk ER. Social class inequality in mortality from 1921 to 1972 in England and Wales. *Popul Studies* 1985;39(1):17-31.
17. Jones IG, Cameron D. Social class analysis—an embarrassment to epidemiology. *J Public Health* 1984;6(1):37-46.
18. Wilkins R. Use of postal codes and addresses in the analysis of health data. *Rapports sur la santé* 1993;5(2):157-77.
19. Wilkins R. *FCCP + Version 5F Guide de l'utilisateur. Logiciel de codage géographique basé sur les fichiers de conversion des codes postaux de Statistique Canada mises à jour en juillet 2009*. Ottawa, Statistique Canada, 2010.
20. Schuurman N, Bell N, Dunn JR, Oliver L. Deprivation indices, population health and geography: An evaluation of the spatial effectiveness of indices at multiple scales. *J Urban Health* 2007;84(4):591-603.
21. Geronimus AT. Invited commentary: Using area-based socioeconomic measures – think conceptually, act cautiously. *Am J Epidemiol* 2006;164(9):835-40.
22. Subramanian SV, Chen JT, Rehkopf DH, et coll. Comparing individual- and area-based socioeconomic measures for the surveillance of health disparities: A multi-level analysis of Massachusetts births, 1989-1991. *Am J Epidemiol* 2006;164(9):823-34.
23. Kreiger N et coll. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? *Am J Epidemiol* 2002;156(5):471-82.
24. Pampalon R, Hamel D, Gamache P. Une comparaison de données socioéconomiques individuelles et géographiques pour la surveillance des inégalités sociales de santé au Canada. *Rapports sur la santé* 2009;20(4):1-12.
25. Macintyre S, Ellaway A, Cummins S. Place effects on health: How can we conceptualise, operationalise and measure them. *Soc Sci Med* 2002;55:125-39.