Socioeconomic Misclassification in Ontario's Health Care Registry

Richard H. Glazier, MD, MPH, CCFP, FCFP[1-4]
Maria I. Creatore, MSc[1]
Mohammad M. Agha, PhD[1]
Leah S. Steele, MD, CCFP[1-3]
for the Inner City Toronto Time Trends Working Group[5]

ABSTRACT

Background: Addresses in some provincial health care registries are not systematically updated. If individuals are attributed to the wrong location, this can lead to errors in health care planning and research. Our purpose was to investigate the accuracy of socioeconomic classification based on addresses in Ontario's provincial health care registry.

Methods: The study setting was Toronto's inner city, an area with a population of 799,595 in 1996. We ordered enumeration areas by 1996 mean household income and divided them into five roughly equal income groups by population. We then assigned an income quintile to each individual using both the address from Ontario's provincial heath care registry and that from hospital discharge abstracts. We compared these two sets of income quintiles and also used them to generate quintile-specific rates of medical hospital admissions in the year 2000.

Results: Provincial registry and hospital-based addresses agreed on the exact enumeration area for 78.1% of individuals and for income quintile for 84.8% of individuals. Disagreement by more than one income quintile occurred for 7.4% of individuals. The two methods of assigning income quintiles yielded income-specific medical hospitalization rates and rate ratios that agreed within 1%.

Interpretation: Although address inaccuracy was found in Ontario's health care registry, serious socioeconomic misclassification occurred at a relatively low rate and did not appear to introduce significant bias in the calculation of hospital rates by socioeconomic group. Updating of addresses at regular intervals is highly desirable and would result in improved accuracy of provincial health care registries.

La traduction du résumé se trouve à la fin de l'article.

1. Inner City Health Research Unit, St. Michael's Hospital, Toronto, ON
2. Department of Family and Community Medicine, University of Toronto
3. Family Healthcare Research Unit, University of Toronto
4. Department of Public Health Sciences, University of Toronto
5. Inner City Toronto Time Trends Working Group members: Mohammad M. Agha, Maria I. Creatore, Richard H. Glazier, Stephen Hwang, Anne Rhodes and Leah Steele from the St. Michael's Hospital Inner City Health Research Unit; Robin Badgley, Jocalyn Clark and Marsha M. Cohen from the Centre for Research in Women's Health, Toronto; Elizabeth M. Badley, Public Health Sciences, University of Toronto; Peter Gozdyra, Department of Geography, University of Toronto; Dianne Patychuk, Toronto Public Health; Lorraine Purdon, Southeast Toronto Project

Correspondence and reprint requests: Dr. Richard Glazier, St. Michael's Hospital, 30 Bond St., Toronto, ON M5B 1W8, Tel: 416-864-5486, Fax: 416-864-5485, E-mail: richard.glazier@utoronto.ca

Supported by: National Health Research and Development Program, Grant #6606-6591-001; Medical Research Council of Canada, Grant #MOP 15693; Institute for Clinical Evaluative Sciences, Toronto; Ontario Ministry of Health and Long-Term Care.

The opinions, results and conclusions are those of the authors and no endorsement by the Ministry is intended or should be inferred.

Socioeconomic factors are important determinants of health status and health care utilization.[1] Measurement of socioeconomic status often takes place at the area level, either to assign socioeconomic status to individuals or for making ecological inferences.[1,2] Most evidence supports a strong relationship between area level socioeconomic characteristics and those of the individual, but this relationship depends on the measures chosen and the population under study.[3-5] Even when individual level data are available, multilevel analyses using area of residence can add valuable information.[6,7] Analyses using area-level data are especially useful and relevant for policy questions that relate to geographic areas, such as public health, hospital, and environmental planning.

Administrative data are increasingly being used for area-level analyses and for that reason, the accuracy of socioeconomic measures in population and health care utilization databases is important. The address information found on health records such as hospital discharge abstracts available from the Canadian Institute for Health Information (CIHI) is relatively reliable for placing people in a geographic location and is updated with each admission.[8] The addresses in some provincial health care registries, however, are not systematically updated. If address information is inaccurate or outdated, the resulting misclassification of area of residence can affect estimates of needs and resource planning and can bias estimates of socioeconomic effects. The purpose of this study was to investigate the accuracy of socioeconomic classification based on addresses in Ontario's provincial health care registry.

METHODS

The study area is located in the central southern portion of the city of Toronto, Ontario, including the city's downtown core, consisting of most of the former city of Toronto, all of the former Borough of East York and a portion of the former City of York. In 1996, it had a population of 799,595 with mean household incomes for enumeration areas (EAs) ranging from $8,622 to $304,454 (median $45,108). EAs were the smallest geographic census units available for analysis in 1996 that included socioeconomic information. On average, our EAs contained 400-600 peo-

ple. Ontario's population-based health care registry for all individuals covered under provincial health care is the Registered Persons' Database (RPDB), which includes each individual's address by 6-digit postal code. Use of the unique Health Number allows linkage of individuals from the RPDB to Ontario's health care utilization databases, including physician claims and drug benefits.

We assigned addresses to EAs using the Statistics Canada Postal Code Conversion File Plus (PCCF+) for 2000. Hospitalized individuals with a study area EA in the CIHI Discharge Abstract Database between April 2000 and March 2001 were linked through their Health Number to their RPDB EA in the same year. Discharges from all Ontario acute care hospitals were included. In the case of multiple hospital admissions by an individual, only the first admission in the fiscal year was included. EAs were ordered by mean household income on the 1996 Canada Census from lowest to highest and divided into quintiles with roughly equal population. EAs for which income was suppressed were excluded from the analysis. In some cases, the RPDB listed more than one address from 1991-2000 due to address changes over time. In the case of multiple addresses in the RPDB, we used the address most closely representing the place of residence in 2000.

In order to demonstrate the effect of socioeconomic misclassification on hospital admission rates, we used medical admissions (using the CIHI Med-Surg indicator) for April 2000 to March 2001 and examined hospital rates for each income quintile using both provincial registry-based and hospital-based income quintiles. We also produced rate ratios for income quintiles compared with quintile 5, the highest income group.

This study received ethics approval from the University of Toronto.

RESULTS

Among eligible inner city residents admitted to hospital from April 2000 to March 2001, the majority of EAs based on addresses found in the provincial registry and hospital discharge abstracts matched exactly (78.1%). Of the remainder, an additional 6.7% were found to be in the same income quintile (Table I). Among EAs in the provincial registry that were in a different quintile from the hospital-based EA, there was an equal split between those in quintiles adjacent (7.8%) and non-adjacent (7.4%) to the hospital quintiles. There was less concordance in income quintiles for individuals who had multiple addresses in the RPDB. Overall, agreement for the same quintile was 84.8% and for the same or an adjacent quintile was 92.6%.

TABLE I

Agreement of Provincial Registry-based Addresses with Hospital-based Addresses, Inner City Toronto, 2000

	Single Addresses (%)	Multiple Addresses* (%)	Total (%)
Total individuals	24,043 (100.0)	17,820 (100.0)	41,863 (100.0)
Same enumeration area	19,704 (82.0)	13,007 (73.0)	32,711 (78.1)
Different enumeration area	4339 (18.0)	4813 (27.0)	9152 (21.9)
Same income quintile	1317 (5.5)	1474 (8.3)	2791 (6.7)
Different income quintile	3022 (12.5)	3339 (18.7)	6361 (15.2)
Adjacent	1572 (6.5)	1699 (9.5)	3271 (7.8)
Non-adjacent	1450 (6.0)	1640 (9.2)	3090 (7.4)

* In the case of multiple addresses in the provincial registry from 1991-2000, the address most closely representing 2000 was used.

Agreement of provincial registry-based and hospital-based EAs by income quintile appears in Table II. Agreement in all quintiles exceeded 80% and was greatest for the lowest income quintile (87.0%) and the highest income quintile (89.3%). In the case of disagreement, there was a slight tendency for the provincial registry-based EAs to be assigned to lower income quintiles than hospital-based EAs.

In Table III, medical admission rates showed high agreement between income quintiles derived from provincial registry-based and hospital-based addresses. These rates differed by less than 1%. Rate ratios for income quintiles compared with quintile 5 (the wealthiest) also showed strong agreement and demonstrated a socioeconomic gradient in medical admissions with the highest rates found in the lowest income quintile.

The hospitalized population differed from the general population of Toronto's inner city in that it was more likely to be female (60.0% versus 51.0%) and older (mean 41.0 versus 36.9 years of age).

DISCUSSION

Lack of routine updating of address information in provincial health registries could affect inferences in health services research and decisions about appropriate allocation of health care. Our findings indicate that for socioeconomic status, the degree of address misclassification in Ontario's provincial health care registry was relatively small and did not result in serious error compared with the use of regularly updated hospital-based addresses. Although we found serious misclassification was infrequent, overall misclassification of income quintile was about 15%, high enough to raise concerns depending on the type of analysis being performed or decision being made. Nevertheless, the impact of socioeconomic misclassification on the calculation of hospital rates and rate ratios was very small and unlikely to introduce serious bias in most types of analyses.

Where misclassification occurred with provincial registry-based addresses, it tended to assign people to a lower socioeconomic group than hospital-based addresses. This effect might be explained by initial location of residents and migrants to Toronto in lower income neighbourhoods, with later movement to better-off neighbourhoods and failure to update the new address in the provincial registry. Those with multiple addresses in the provincial registry were also more likely to represent a mobile population, so it is not surprising that agreement with hospital-based location would be lower for this group.

The study was based on individuals admitted to hospital for whom a valid health number and valid address could be found in the hospital discharge abstract. Individuals without health coverage, homeless individuals and those who may have appeared in the provincial health care registry but not the hospital abstracts were not included in this study. Inner city residents not admitted to hospital and therefore not included in the study were younger and more likely to be male than those in the study. It is possible that socioeconomic mobility among those not

TABLE II

Comparison of Income Quintiles Derived from Provincial Registry-based and Hospital-based Addresses, Inner City Toronto, 2000

Provincial Registry-based Income Quintiles	Hospital-based Income Quintiles					
	1	2	3	4	5	Total
1	8430 (87.0%)	476 (5.8%)	344 (4.2%)	303 (3.7%)	117 (1.5%)	9670
2	425 (4.4%)	6621 (81.2%)	466 (5.8%)	329 (4.0%)	173 (2.2%)	8014
3	358 (3.7%)	510 (6.2%)	6676 (82.6%)	389 (4.8%)	208 (2.7%)	8141
4	321 (3.3%)	347 (4.2%)	368 (4.6%)	6835 (83.8%)	334 (4.3%)	8205
5	156 (1.6%)	202 (2.5%)	232 (2.9%)	303 (3.7%)	6940 (89.3%)	7833
Total	9690 (100%)	8156 (100%)	8086 (100%)	8159 (100%)	7772 (100%)	41,863

TABLE III

Medical Hospitalization Rates per 1000 Population and Rate Ratios* by Income Quintile for Provincial Registry-based and Hospital-based Addresses, Inner City Toronto, 2000

	Hospital-based Quintiles				Provincial Registry-based Quintiles			
Income Quintile	Individuals Admitted	Population	Rate	Rate Ratio	Individuals Admitted	Population	Rate	Rate Ratio
1	2723	117,930	23.1	1.23	2796	117,930	23.7	1.25
2	2441	120,765	20.2	1.07	2508	120,765	20.8	1.10
3	2516	121,470	20.7	1.10	2486	121,470	20.5	1.08
4	2650	124,485	21.3	1.13	2665	124,485	21.4	1.13
5	2271	120,690	18.8	1.00	2286	120,690	18.9	1.00
Total	12,601	605,340	20.8		12,741	605,340	21.0	

* Rate ratios compare income quintile rates with the rate of income quintile 5 (wealthiest).

admitted was higher than among those admitted. In that case, our study of hospitalized individuals would underestimate the degree of socioeconomic misclassification in the provincial registry.

These results should be applied to other settings with caution. Our study may have overestimated the degree of misclassification in provincial health care registries by using very small geographic areas in a setting with high mobility. Misclassification would be expected to be lower for larger geographic areas such as census tracts and forward sortation areas and for geographic areas with lower mobility than inner city Toronto. For other settings, however, misclassification could be larger than we found in inner city Toronto. Rural areas, for example, have scattered populations, shared post office boxes and more problematic conversion of postal to census geography. In other provinces where the health care registry is not routinely updated, rurality, patterns of population mobility and the lag time in address updating need to be taken into account in applying these findings.

Various mechanisms could be used to ensure routine updating of addresses in provincial health care registries. Physicians who bill provincial plans could submit an updated address with each claim. These are likely to be accurate since professional practice dictates keeping current contact information for each patient. Extra costs for billings and potential infringement of patient privacy are among the barriers to this approach. Routine updating of health cards and/or health numbers would be another approach. This has been partially implemented in Ontario where some but not all health cards carry expiry dates. Leaving the onus on members of the public to inform a ministry of health when they move is unlikely to result in comprehensive updating.

We conclude that socioeconomic misclassification exists in Ontario's health care registry due to lack of updated addresses. Serious misclassification, however, occurs at a relatively low rate and does not appear to introduce significant bias in the calculation of hospital rates by socioeconomic group. Updating of addresses at regular intervals is highly desirable and would result in improved accuracy of provincial health care registries.

…continues next page

RÉSUMÉ

Contexte : Les adresses de certains registres provinciaux de la santé ne sont pas systématiquement mises à jour. Or, des adresses erronées peuvent entraîner des erreurs dans la planification et la recherche en matière de soins de santé. Notre étude porte sur l'exactitude de la classification socio-économique fondée sur l'adresse dans le registre provincial de la santé de l'Ontario.

Méthode : L'étude a porté sur les quartiers déshérités du centre-ville de Toronto, qui comptaient 799 595 habitants en 1996. Nous avons classé les secteurs de dénombrement selon le revenu moyen des ménages en 1996, et nous les avons divisés en cinq catégories de revenu à peu près égales, par population. Nous avons ensuite affecté un quintile de revenu à chaque personne, selon son adresse dans le registre provincial de la santé de l'Ontario et les registres des sorties des hôpitaux. Nous avons comparé ces deux ensembles de quintiles de revenu, que nous avons utilisés pour produire des taux par quintile pour les malades reçus dans les hôpitaux en 2000.

Résultats : Le secteur de dénombrement figurant dans le registre provincial et dans les adresses obtenues des hôpitaux concordait dans 78,1 % des cas, et le quintile de revenu, dans 84,8 % des cas. Dans 7,4 % des cas, l'écart était supérieur à un quintile. Les deux méthodes d'affectation du quintile de revenu ont donné des taux d'hospitalisation par revenu et des ratios des taux qui concordaient dans une limite de 1 %.

Interprétation : Malgré les erreurs d'adresse trouvées dans le registre de la santé de l'Ontario, les erreurs de classification socio-économique graves sont relativement rares et ne semblent pas induire un biais important dans le calcul des taux d'hospitalisation par groupe socio-économique. La mise à jour périodique des adresses serait toutefois hautement souhaitable et contribuerait à améliorer l'exactitude des registres provinciaux de la santé.

REFERENCES

1. Roos NP, Mustard CA. Variation in health and health care use by socioeconomic status in Winnipeg, Canada: Does the system work well? Yes and no. *Millbank Q* 1997;75:89-111.
2. Alter DA, Naylor CD, Austin P, Tu JV. Effects of socioeconomic status on access to invasive cardiac procedures and on mortality after acute myocardial infarction. *N Engl J Med* 1999;341:1359-67.
3. Krieger N. Overcoming the absence of socioeconomic data in medical records: Validation and application of a census based methodology. *Am J Public Health* 1992;92:703-10.
4. Demissie K, Hanley JA, Menzies D, Joseph L, Ernst P. Agreement in measuring socio-economic status: Area-based versus individual measures. *Chron Dis Can* 2000;21:1-7.
5. Mustard CA, Derksen S, Berthelot JM, Wolfson M. Assessing ecologic proxies for household income: A comparison of household and neighbourhood level income measures in the study of population health status. *Health Place* 1999;5(2):157-71.
6. Kennedy BP, Kawachi I, Glass R, Prothrow-Stith D. Income distribution, socioeconomic status, and self rated health in the United States: Multilevel analysis. *BMJ* 1998;317:917-21.
7. Duncan C, Jones K, Moon G. Context, composition and heterogeneity: Using multilevel models in health research. *Soc Sci Med* 1998;46:97-117.
8. Appendix. A Summary of Studies on the Quality of Health Care Administrative Databases in Canada, p 341-342. In: Goel V, Williams JI, Anderson GM, Blackstien-Hirsch, Fooks C, Naylor CD (Eds.), *Patterns of Health Care in Ontario. The ICES Practice Atlas,* 2nd edition. Ottawa: Canadian Medical Association, 1996.

Received: November 5, 2001
Accepted: September 12, 2002

Directives de rédaction de la RCSP

La *Revue canadienne de santé publique* publie des articles originaux sur tous les aspects de la santé publique, de la médecine préventive et de la promotion de la santé qui ont été évalués par des pairs. Tous les manuscrits soumis pour publication dans ses colonnes doivent respecter les Directives de rédaction de la RCSP. On trouvera une version complète sur le site Web de l'ACSP à l'adresse suivante < www.cpha.ca/francais/cjph/stylreq/style.htm >. Les auteurs qui n'ont pas accès à Internet peuvent se reporter aux pages 7-8 du numéro de janvier/février 2002 (Vol. 93, No. 1) de la *Revue canadienne de santé publique* ou se mettre en rapport avec le bureau de la rédaction pour en recevoir un exemplaire par télécopieur.

La Rédaction doit recevoir un exemplaire imprimé de l'original et deux copies pour l'évaluation par les pairs (dactylographié à double interligne, sur un seul côté de la feuille) avec toutes les figures et tous les tableaux, et accompagné d'une version électronique. Veuillez numéroter les pages les unes à la suite des autres, y compris le résumé et sa traduction (le cas échéant), le texte, les remerciements, les références, les tableaux et les figures.

Les articles originaux ne doivent pas dépasser **2 000 mots**. Veuillez indiquer le nombre de mots de votre article. Les brefs comptes-rendus peuvent être soumis et doivent inclure des résumés structurés (voir le paragraphe « Les documents sont… »)

La Revue a adopté une politique d'évaluation des articles par des pairs, en double aveugle. Auteurs et évaluateurs restent anonymes. Afin d'assurer l'anonymat de cette évaluation par des pairs, les auteurs doivent s'assurer de ne donner les renseignements les identifiant que sur la page titre de **l'original seulement**; seul le titre doit apparaître sur les deux copies à l'intention des évaluateurs.

La page titre de l'original doit comporter : 1) le titre; 2) un bref sous-titre de 40 caractères au maximum (y compris les blancs); 3) les noms, prénoms et initiales du milieu (le cas échéant) des auteurs, accompagnés de leurs diplômes d'études de plus haut niveau avec les noms du service et de l'organisme dans le cadre duquel le travail publié a été effectué; 4) le nom, l'adresse, le numéro de téléphone, le numéro de télécopieur et l'adresse électronique de l'auteur à qui adresser toute correspondance; 5) le nom et l'adresse de l'auteur à qui les demandes de réimpression doivent être envoyées (s'il s'agit de la même personne, regroupez 4 et 5 ensemble).

Les déclarations relatives aux limites de responsabilité et les mentions concernant l'aide reçue doivent aussi figurer sur la page titre. Une lettre signée par tous les auteurs comme quoi ils acceptent d'être publiés doit être incluse.

Les documents sont acceptés en anglais ou en français. La page deux doit présenter un résumé structuré pouvant avoir jusqu'à 250 mots, avec les titres : Objectif, Méthodes, Résultats, Discussion (ou Interprétation). En outre, la *RCSP* demande une traduction professionnelle du résumé. Pour garantir une qualité constante de la traduction de la Revue, la *RCSP* se réserve le droit de faire retraduire les résumés soumis, aux frais des auteurs après les en avoir avisés. S'ils le préfèrent, les auteurs peuvent demander de faire traduire directement leurs résumés par le traducteur officiel de la Revue à raison de 22 cents le mot plus la TPS de 7 % (ou le TVH de 15 %).

Nous vous encourageons à écrire à la Rédaction, en vous priant d'envoyer des lettres aussi courtes que possible.

La Rédaction se réserve le droit de modifier les articles.

Les documents à publier doivent être soumis au rédacteur scientifique de la Revue canadienne de santé publique, 1565 avenue Carling, bureau 400, Ottawa, Ontario (Canada) K1Z 8R1.