

Prognostic Relevance of Census-derived Individual Respondent Incomes Versus Household Incomes

Danielle A. Southern, MSc^{1,2}

Peter D. Faris, PhD³

Merril L. Knudtson, MD²

William A. Ghali, MD, MPH^{1,2,4}

for the APPROACH Investigators

ABSTRACT

Background: Census-based measures of income derived from median income of a geographic area are often used in health research. Many national census surveys gather information on both the respondent's individual income and the income for the entire household, giving researchers a choice of census income measures. We compared the extent to which individual respondent income and household income (both obtained from census data) are associated with outcomes in a cohort of patients with cardiac disease.

Methods: We used data from the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease (APPROACH), where postal codes were linked to the Postal Code Conversion File (PCCF) to determine each patient's census Dissemination Areas (DA). DA-derived median household income and median individual income were obtained from the 2001 Canadian Census and survival outcomes were then directly determined for income groupings defined by quintile. Two-year survival adjusted for age and sex was described with a proportional hazards analysis.

Results: There were 9,397 patients undergoing cardiac catheterization between January 1, 2001 and March 31, 2002, with complete DA-level median income measures. Household income quintiles yielded a wider spread of survival across quintiles (range of 2-year estimated survival, 91.8% to 95.9% for household income versus 92.8% to 95.6% for respondent income), as well as a more progressive decline in survival as income decreased. This progressive decline was not seen for the respondent income measure.

Conclusions: The greater spread and progressive decline of survival for household income relative to respondent income leads us to conclude that household income is the better socio-economic determinant of health in our data and for the outcome measure we studied.

MeSH terms: Censuses; socioeconomic status; income; survival analysis; registries

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

University of Calgary, Calgary, AB

1. Department of Community Health Sciences

2. Department of Cardiac Sciences

3. Centre for Health and Policy Studies

4. Department of Medicine

Correspondence and reprint requests: Dr. William A. Ghali, HSG239, 3330 Hospital Dr. NW, Calgary, AB T2N 4N1, Tel: 403-210-9317, Fax: 403-210-3818, E-mail: wghali@ucalgary.ca

Acknowledgements: APPROACH Clinical Steering Committee: Edmonton: S. Archer, M.M. Graham, W. Hui (Chair), A. Koshal, and R.T. Tsuyuki. Calgary: L.B. Mitchell, M. Traboulsi, W.A. Ghali, M.L. Knudtson and A. Maitland.

Financial support: APPROACH was funded in 1995 by the WG Weston Foundation, with ongoing support from Merck Frosst Canada Inc, Monsanto Canada Inc-Searle, Eli Lilly Canada Inc, Guidant Corporation, Boston Scientific Ltd, Hoffmann-La Roche Ltd, Johnson & Johnson Inc-Cordis, and the Province-Wide Services Committee of Alberta Health and Wellness.

Dr. Ghali is supported by a Health Scholar Award from the Alberta Heritage Foundation for Medical Research, and by a Government of Canada Research Chair in Health Services Research.

Many national census surveys¹⁻³ gather information on both the respondent's income and the income for the entire household. Several studies in the literature have used household income as a marker of socio-economic status (SES).⁴⁻¹⁷ Others, however, have chosen to use personal median income (also called individual or respondent median income) as the main income measure.¹⁸⁻²⁰ The question as to whether household or respondent median income is a better measure of SES as a determinant of health outcomes is therefore very pertinent. Somewhat surprisingly, however, there is a lack of information comparing the prognostic relevance of individual respondent incomes and household incomes.

The objective of our paper is to guide the choice of income measure by comparing the extent to which individual respondent income and household income (both from census data) are associated with survival outcomes in a cohort of patients with cardiac disease.

METHODS

Data sources

The study population was derived from the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease (APPROACH),²¹ an inception cohort database that includes all patients in Alberta, Canada undergoing cardiac catheterization. Patients are followed longitudinally after cardiac catheterization until end of follow-up or death. The latter is ascertained through semi-annual linkage to death records from the Alberta Bureau of Vital Statistics. We analyzed data on patients registered in APPROACH from January 1, 2001 to March 31, 2002. Patients were excluded if they were missing a valid postal code or were not Alberta residents.

Information on household income

Statistics Canada census data from 2001 were used as a source of respondent median income (median individual income in Statistics Canada dictionary) and median household income for each dissemination area (DA), of which Alberta has a total of 5,143.²² A previous paper has shown that enumeration area (now referred to as the DA) is a better proxy for income than forward sortation area (FSA),²³ and so for this

reason, we used census data at the DA level for our current study.

The Postal Code Conversion File (PCCF) for January 2003²⁴ contains all 826,709 postal codes ever used by Canada Post Corporation since 1983 (including many that are now retired). Each postal code in this file is linked to one or more DAs. When there is more than one DA for a postal code, Statistics Canada provides a single link indicator (SLI) to select the most representative DA. After merging the APPROACH data with the PCCF, using only the SLI, we merged the new file with the census data files containing DA respondent median income and household median income.

Application of household and respondent median income categories in outcome analyses

Both respondent and household median incomes were then broken down into quintiles based on the distribution of DA-derived incomes seen in our study population. We assessed survival across income quintiles for each of the income measures and presented this graphically using Kaplan-Meier plots. We also performed an adjusted survival analysis using proportional hazards models that described survival risk by quintile, controlling for age and sex using the corrected group prognosis method.²⁵ Comorbidity and severity of disease are available in the APPROACH registry. We did not, however, include them in the proportional hazards analysis because such variables are potential mediators of the association between income and survival – an association that we did not want to attenuate by controlling for variables that mechanistically may contribute to the association.

Analyses were performed using SAS V8.1. APPROACH is approved by the ethics committees at the Universities of Calgary and Alberta.

RESULTS

Characteristics of study sample

Our starting point for analysis was an APPROACH analysis file containing 10,817 patients; 10,627 with complete postal code information and 9,818 of these were Alberta residents. Following linkage to DA-based income measures, 9,392 of these patients had information on respon-

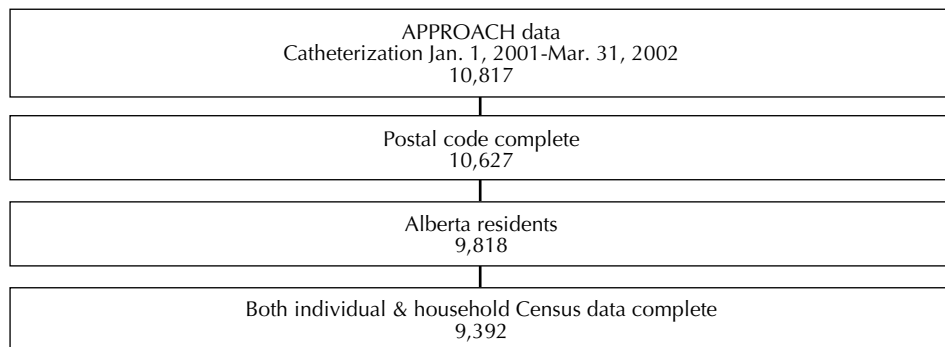


Figure 1. Derivation of study population

TABLE I
Characteristics of Study Sample

Characteristics	N=9,392
Age (mean years (Std.Dev.))	62.6 (12.0)
Male (N (%))	6,423 (68.4%)
Respondent Income	
Mean (Std.Dev.)	\$23,983 (6,485)
Median (IQR)	\$23,086 (8,776)
Household Income	
Mean (Std.Dev.)	\$56,207 (21,988)
Median (IQR)	\$52,884 (27,344)
Respondent Income Quintiles	
Lowest	\$5,264-18,460 (N=1,880)
Second	\$18,467-21,382 (N=1,879)
Third	\$21,388-25,093 (N=1,878)
Fourth	\$25,095-29,439 (N=1,878)
Highest	\$29,447-70,961 (N=1,877)
Household Income Quintiles	
Lowest	\$11,596-38,304 (N=1,882)
Second	\$38,326-48,259 (N=1,879)
Third	\$48,268-58,236 (N=1,878)
Fourth	\$58,270-72,462 (N=1,885)
Highest	\$72,484-221,910 (N=1,868)

TABLE II
Hazard Ratios (HR) for Death After Cardiac Catheterization

Income Quintiles	Crude HR (95% CI)	Adjusted* HR (95% CI)
Respondent Income		
Lowest	1.51 (1.14, 2.00)	1.44 (1.08, 1.90)
Second	1.48 (1.11, 1.96)	1.44 (1.09, 1.92)
Third	1.44 (1.08, 1.92)	1.40 (1.05, 1.86)
Fourth	1.38 (1.03, 1.84)	1.36 (1.02, 1.82)
Highest	1.00	1.00
Household Income		
Lowest	1.90 (1.43, 2.52)	1.68 (1.27, 2.23)
Second	1.56 (1.17, 2.09)	1.45 (1.08, 1.94)
Third	1.64 (1.23, 2.19)	1.52 (1.14, 2.03)
Fourth	1.25 (0.92, 1.69)	1.22 (0.90, 1.66)
Highest	1.00	1.00

* Adjusted for age and sex

dent and household median income available from the census linkage (Figure 1). The demographic characteristics (age, sex, range of income in quintiles) of our sample are presented in Table I.

Application to analyses of two-year survival

The relationship of income categories derived from respondent and household

median income to survival differs across income measures. Figure 2 presents crude Kaplan-Meier plots for survival extending to two years by income category derived from respondent median income (Panel A) and household median income (Panel B). These analyses reveal that both income measures provide prognostic information, with progressively poorer survival in lower income groupings. However, the relation-

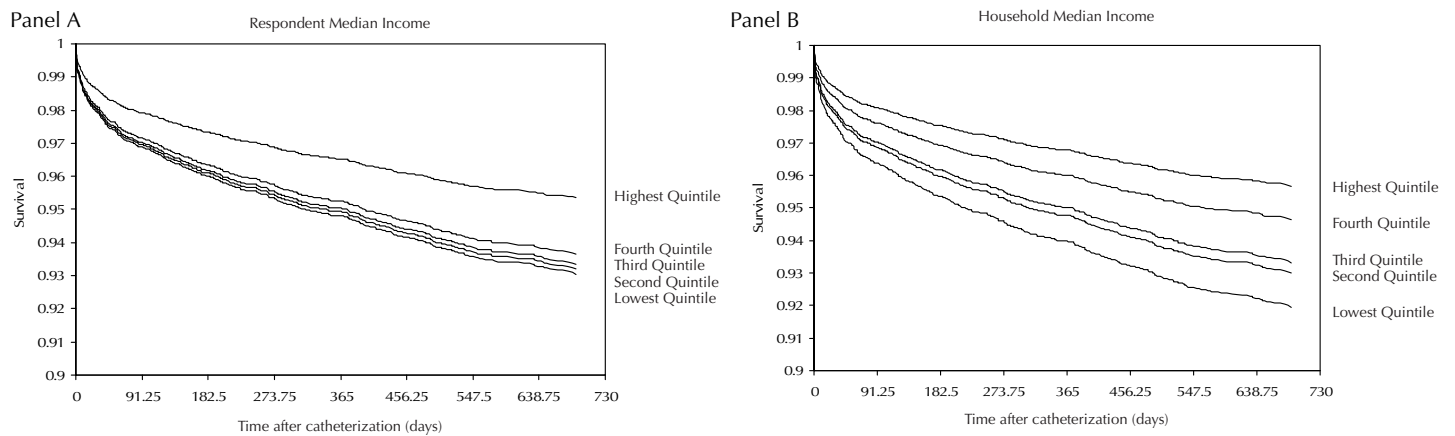


Figure 2. Crude respondent level survival curve (PANEL A) and crude household level survival curve (PANEL B)

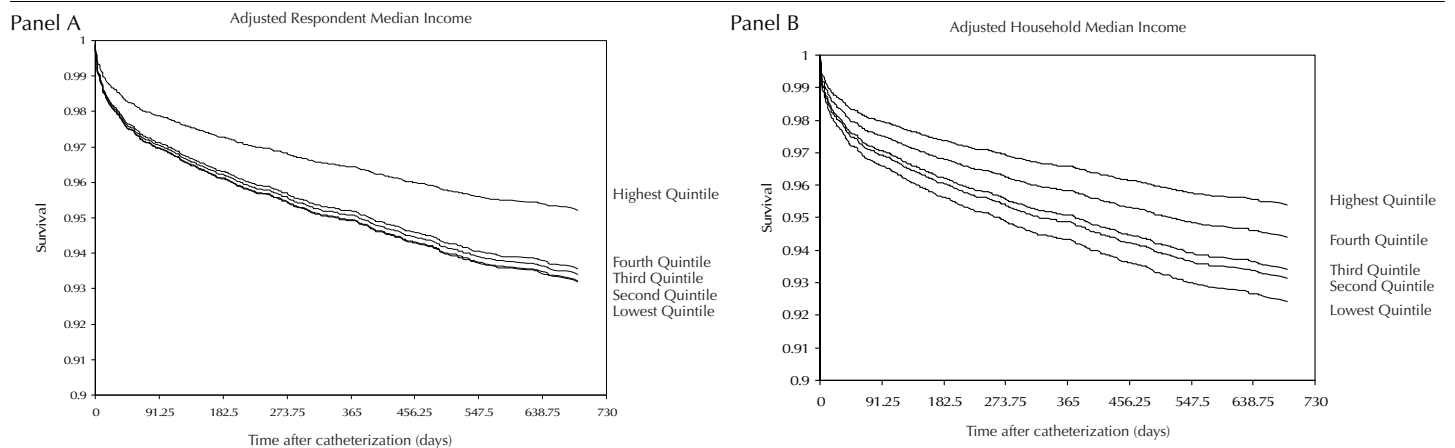


Figure 3. Adjusted respondent level survival curve (PANEL A) and adjusted household level survival curve (PANEL B)

ship is more clearly demarcated and progressively consistent for household median income than it is for respondent median income. Both crude and adjusted hazard ratios for death following cardiac catheterization, grouped by household and respondent incomes, are presented in Table II. The adjusted hazard ratios have been adjusted for age and sex. There is an increase in hazard as income decreases. The range of respondent median income hazard ratios by quintile is tightly clustered (1.34-1.41), whereas the range of household median income hazard ratios by quintile is more widely spread (1.22-1.63). The corresponding adjusted survival curves are presented in Figure 3.

DISCUSSION

Our findings demonstrate that both respondent incomes and household incomes are prognostically relevant in a population of cardiac patients. However, the relationship between household median income and survival is more consistent

and steeply graded than that of respondent median income. These findings suggest that household median income may better represent socio-economic status as a determinant of health outcomes than does individual median income.

Studies using household income demonstrate an income-outcome gradient.^{7,9} Studies using respondent income also find an income-outcome gradient.¹⁸⁻²⁰ Our study therefore confirms previous findings from all these prior studies indicating that census-derived income measures are prognostically relevant, but expands on these findings by demonstrating that household income may be a more meaningful and discriminating measure.

The essence of a 'household' is a sharing of resources and/or capacity among individuals toward a common good.²⁶ Bonney²⁶ presents the assumption that all members of a household unit are potential labour market participants whose actions may have significant influences on the social class position and life chances of the unit and its constituent members. The

term 'household' is also used rather than 'family' to allow for a wider variety of actual domestic arrangements for performing household labour.²⁶ In concrete terms, the income of a stay-at-home partner is not necessarily relevant if the other partner is a high earner. However, there is also a need to acknowledge that aggregate measures of household income can be misleading if the typical number of individuals per household is unusually large in certain areas – a situation that could result in fewer economic resources per individual within households. If we extend the analysis presented in Figure 3, Panel B, to account for household size, by dividing household income by the mean number of residents in a household (as recorded in census data by dissemination area), we find an attenuated separation of the five survival curves defined by quintile (with hazard ratios ranging from only 1.0 [reference] to 1.28). We suspect that this occurs because high numbers of dependants (e.g., children) per household can dilute the significance of income, as there presumably comes a point

where it becomes misleading to attribute equal share of household income to a 3rd, 4th, or even a 5th dependant. These caveats aside, however, our findings indicate that household income is more informative as a predictor of health outcomes than are both individual respondent income and household income divided by household size.

Some caveats and limitations should be noted. First, we only assessed the single outcome of mortality. An important caveat is that household income does not provide insight into some relevant issues such as gender equity in incomes. Also, readers should note that true individual income is not assessed here. However, area-based incomes derived from census have nonetheless been shown to be important. A final caveat is that we are not necessarily able to generalize our findings to other databases and/or other outcomes. On this latter point, we encourage others to explore whether our findings are replicable in other data sources.

Despite these caveats, our results are important for researchers using data that rely on census-based measures of income to assess socio-economic status. Recognizing the final caveat mentioned above, we recommend that health researchers using census-based income measures should assess both respondent and household income and select the most appropriate measure for their situation. In our data, the greater spread and progressive decline of survival for household income relative to respondent income leads us to conclude that household income better represents socio-economic status as a determinant of health.

REFERENCES

1. Statistics Canada. 2001 Census Dictionary. Ottawa, ON. 2003; Cat 92-378-XIE.
2. US Bureau of Census. Census 2000. Washington, DC: US Bureau of Census.
3. UK Census. Census 2001. London, England: National Statistics.
4. Kephart G, Thomas VS, MacLean DR. Socioeconomic differences in the use of physician services in Nova Scotia. *Am J Public Health* 1998;88(5):800-3.
5. Blanchard JF, Moses S, Greenaway C, Orr P, Hammond GW, Brunham RC. The evolving epidemiology of chlamydial and gonococcal infections in response to control programs in Winnipeg, Canada. *Am J Public Health* 1998;88(10):1496-502.
6. Cairney J, Arnold R. Socioeconomic position, lifestyle and health among Canadians aged 18 to 64: A multi-condition approach. *Can J Public Health* 1998;89(3):208-12.
7. Mustard CA, Roos NP. Variation in health care use by socioeconomic status in Winnipeg, Canada: Does the system work well? Yes and no. *Milbank Q* 1997;75(1):89-111.
8. Diez-Roux AV, Merkin SS, Arnett D, Chambless L, Massing M, Nieto FJ, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med* 2001;345(2):99-106.
9. Veenstra G. Economy, community and mortality in British Columbia, Canada. *Soc Sci Med* 2003;56:1807-16.
10. Wolfson M. Income inequality and mortality among working-age people in Canada and the US. *Health Reports* 1999;11(3):77-82.
11. Galea S, Ahern J, Vlahov D, Coffin PO, Fuller C, Leon AC, Tardiff K. Income distribution and risk of fatal drug overdose in New York City neighborhoods. *Drug Alcohol Depend* 2003;70:139-48.
12. Wen M, Browning CR, Cagney KA. Poverty, affluence, and income inequality: Neighborhood economic structure and its implications for health. *Soc Sci Med* 2003;57:843-60.
13. Wilson K, Elliott S, Law M, Eyles J, Jerrett M, Keller-Olaman S. Linking perceptions of neighborhood to health in Hamilton, Canada. *J Epidemiol Community Health* 2004;58:192-98.
14. Lipton RB, Drum M, Li S, Choi H. Social environment and year of birth influence type 1 diabetes risk for African-American and Latino children. *Diabetes Care* 1999;22(1):78-85.
15. Pomerantz WJ, Dowd MD, Buncher CR. Relationship between socioeconomic factors and severe childhood injuries. *J Urban Health* 2007;8(1):141-51.
16. Kaufman JS, Dole N, Savitz DA, Herring AH. Modeling community-level effects on preterm birth. *AEP* 2003;13(5):377-84.
17. Sin DD, Svenson LW, Man SFP. Do area-based markers of poverty accurately measure personal poverty? *Can J Public Health* 2001;92(3):184-87.
18. 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(18):1359-67.
19. Alter DA, Austin P, Naylor CD, Tu JV. Factoring socioeconomic status into cardiac performance profiling for hospitals. *Med Care* 2002;40(1):60-67.
20. Alter DA, Naylor CD, Austin P, Chan BTB, Tu JV. Geography and service supply do not explain socioeconomic gradients in angiography use after acute myocardial infarction. *CMAJ* 2003;168(3):261-64.
21. Ghali WA, Knudtson ML, on behalf of the APPROACH Investigators. Overview of the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease. *Can J Cardiol* 2000;16(10):1225-30.
22. Statistics Canada. Census of Canada 2001. Alberta info [computer files].
23. Southern DA, Ghali WA, Faris PD, Norris CM, Galbraith PD, Graham MM, Knudtson ML. Misclassification of income quintiles derived from area-based measures: A comparison of enumeration area and forward sortation area. *Can J Public Health* 2002;93(6):465-69.
24. Statistics Canada (Geography Division). Postal Code Conversion File. January 2003 version. Ottawa, ON: 2003.
25. Ghali WA, Quan H, Brant R, van Melle G, Norris CM, Faris PD, et al. Comparison of 2 methods for calculating adjusted survival curves from proportional hazards models. *JAMA* 2001;286:1494-97.
26. Bonney N. Gender, household and social class. *Br J Social* 1988;39(1):28-46.

Received: November 10, 2004

Accepted: May 30, 2005

RÉSUMÉ

Contexte : Les mesures basées sur le recensement du revenu dérivé du revenu médian d'un secteur géographique sont souvent employées dans la recherche de la santé. Beaucoup d'enquêtes nationales de recensement recueillent l'information sur le revenu individuel du répondant ainsi que le revenu pour le ménage entier, donnant aux chercheurs un choix de mesures de revenu. Nous avons comparé le point auquel le revenu individuel du répondant et le revenu du ménage (tous les deux obtenus à partir de données de recensement) sont associés aux résultats dans une cohorte de patients présentant pour une cathérisation cardiaque.

Méthodes : Nous avons employé des données du projet Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease (APPROACH), où des codes postaux ont été liés au fichier de conversion des codes postaux plus (FCCP+) pour déterminer les aires de diffusion (AD) du recensement de chaque patient. Le revenu du ménage médian dérivé des AD et le revenu individuel médian ont été obtenus à partir du recensement du Canada de 2001, et des résultats de survie ont été directement déterminés pour des groupements de revenu définis par quintile. La survie de deux ans ajustée à l'âge et au sexe a été décrite avec une analyse de risques proportionnels.

Résultats : Il y avait 9 397 patients subissant la cathérisation cardiaque entre le 1 janvier, 2001 et le 31 mars, 2002, avec des mesures médianes de revenu de niveau des AD complets. Les quintiles de revenu du ménage ont rapporté une diffusion plus large de survie à travers des quintiles (tranche de survie de 2 ans estimé varie de 91,8 % à 95,9 % pour le ménage, et de 92,8 % à 95,6 % pour le répondant), aussi bien qu'un déclin plus progressif dans la survie pendant que le revenu diminue. Ce même déclin n'a pas été vu pour la mesure de revenu du répondant.

Interprétation : La diffusion plus grande et le déclin progressif de la survie pour le revenu du ménage par rapport à celui du répondant nous mènent à conclure que le revenu du ménage représente mieux le statut socio-économique comme déterminant de la santé dans nos données, et pour le résultat que nous avons évalué.