Learning From the Census: The Socio-economic Factor Index (SEFI) and Health Outcomes in Manitoba

Dan Chateau, PhD, Colleen Metge, PhD, Heather Prior, MSc, Ruth-Ann Soodeen, MSc

ABSTRACT

Objectives: Using data from the Canadian census, researchers at the Manitoba Centre for Health Policy sought to create an area-based socio-economic measure (ABSM). The degree of association between the ABSM and health was evaluated.

Methods: Values on several census variables (including income, education, employment and family structure) were captured at the enumeration-area or dissemination-area level and submitted to a principal components factor analysis to create three ABSMs: an updated version of the Socio-economic Factor Index (SEFI-2) and modified versions of Pampalon’s material deprivation and social deprivation indices. Factor scores from these analyses were then compared with several population health measures: Premature Mortality Rate (PMR), Potential Years of Life Lost (PYLL), life expectancy, and self-rated health.

Results: SEFI-2 scores were strongly related not only to the other ABSMs but also to every measure of health status. The strongest correlations between an ABSM and health measure were for SEFI-2 and PYLL (r=0.85), and SEFI-2 and PMR (r=0.80). The weakest correlations were found with the social deprivation ABSM measure and the self-rated health measure.

Conclusions: ABSMs based on measures from the Canadian census are a valuable resource to population health researchers. Importantly, depending on the research question and reason for the inclusion of an ABSM, these composite measures may perform better than a simple measure of income alone. The ability to adjust for socio-economic status when assessing population health status or population health interventions contributes to the validity of conclusions drawn when conducting this type of research, and ABSMs may be able to substitute for area health status where it may not be easily determined.

Key words: Socio-economic status; socio-economic factors; mortality, premature; censuses; health status

© Canadian Public Health Association, 2012. All rights reserved.

The Canadian census and the Socio-economic Factor Index (SEFI)

In epidemiological or population health research, the means of measuring socio-economic status (SES) has become an important issue. Gathering data from individuals can be expensive and impractical, particularly for organizations involved in population health research, such as the Manitoba Centre for Health Policy (MCHP). As an alternative to this, the Canadian census has been most useful in providing a rich array of information on people living in a relatively small area. In the 2006 census, there were approximately 2,100 dissemination areas in Manitoba; with an average population of about 700 persons, this is the smallest area at which Statistics Canada can report data. During the past two decades, MCHP has been using census data to create several different area-based socio-economic measures (ABSMs). We will describe the development of the Socio-economic Factor Index (SEFI and SEFI-2), as well as measures of deprivation, and compare these ABSMs to several population health outcomes.

The Canadian census and the Socio-economic Factor Index (SEFI)

Average household income is the simplest ABSM to take from the census. Although it may be directly related only to the “economic” aspect of SES, it is conceptually easy to understand and explain, and therefore tempting to use on its own. However, it has a major limitation: information on income is subject to the most stringent criteria for dissemination by Statistics Canada, requiring a minimum area population of 250 or more in order to be reported. Before 2001, the census was envisioned more strictly as a counting mechanism, and the choice of language and boundaries reflected that fact. The smallest area for which data were reported was the enumeration area (EA) rather than the dissemination area (DA), and EA populations were more commonly below the minimum required for reporting.

In the 1996 census, 607 of 2,050 EAs (29.6%) in Manitoba had suppressed values. This improved in the 2001 census, when DAs were adopted, and only 448 of 2,235 DAs (20.0%) were suppressed. By 2006, the spirit of dissemination had truly taken root in the Canadian census, when only 131 of 2,152 DAs (6.1%) in Manitoba had suppressed values. Thus, until 2006, researchers could not rely simply on income or on any ABSM including income.

© Canadian Public Health Association, 2012. All rights reserved.

Author Affiliations
1. Research Scientist, Manitoba Centre for Health Policy, University of Manitoba, Winnipeg, MB
2. Director, Research & Evaluation Unit, Research & Applied Learning Division, Winnipeg Regional Health Authority, Winnipeg, MB
3. Data Analyst, Manitoba Centre for Health Policy, University of Manitoba, Winnipeg, MB
4. Lead Research Coordinator, Manitoba Centre for Health Policy, University of Manitoba, Winnipeg, MB

Correspondence: Dan Chateau, Manitoba Centre for Health Policy, University of Manitoba, 408-727 McDermot Ave., Winnipeg, MB R3E 3P5, Tel: 204-975-7767, Fax: 204-789-3910, E-mail: dan_chateau@cpe.umanitoba.ca

Acknowledgements: This work was supported through funding provided by the Department of Health of the Province of Manitoba to the University of Manitoba (HIPC # 2005/2006-06). The results and conclusions are those of the authors, and no official endorsement by Manitoba Health was intended or should be inferred. Data used in this study are from the Population Health Research Data Repository housed at the Manitoba Centre for Health Policy, University of Manitoba, and were derived from data provided by Manitoba Health.

Conflict of Interest: None to declare.

© Canadian Public Health Association, 2012. All rights reserved.
To overcome the obstacle of suppression of income data, researchers at MCHP instead developed a measure that did not include income. In those previous censuses (and even now) Statistics Canada applied a much less stringent criterion for suppression of other types of information, such as education and employment. Included in this first version of the Socio-economic Factor Index (SEFI) were high school graduation rates, unemployment rates, proportion of single-parent families in an area, female labour force participation, and an area's age dependency ratio. A principal components factor analysis was conducted on these variables.

The basic premise of factor analysis is fairly straightforward: variables that are correlated are grouped together. That is, when comparing the different variables that make up an index (e.g., average income and high school graduation rates), one might find relations (e.g., as high school graduation rate increases, income increases) that reflect the character of SES. Factor analysis looks for this common variance among the variables and creates a mathematical combination of the included variables, in which a certain portion of each indicator contributes to an overall factor score. The resulting factor scores were the ABSM values used in subsequent analyses. The contribution of a variable to the factor is known as a factor loading and can range from -1 to +1. The larger the absolute size of the loading, the more important that variable is for the factor. Initially, these analyses were all conducted at larger area levels, such as regional health authorities, where the census variables and the health outcomes to which they were being compared (e.g., hospitalization rates, physician visit rates, mortality rates) could be summarized.

With the adoption of the DA by Statistics Canada in 2001, the inclusion of income as a direct measure of economic status in the ABSM was possible. This new version (SEFI-2) included four variables from the census, analyzed at the DA level: average household income, proportion of high school graduates, unemployment rate, and proportion of single-parent families. The factor scores for this new version are presented in Table 1. The strong relation between the variables is evident in the high loadings, particularly for income and education. It is of note that, as a risk score, higher SEFI-2 values represent lower SES, and lower SEFI-2 values represent higher SES. From this analysis, it is apparent that if one were to choose a single variable to represent SES, income would be the logical choice, as it has the highest factor loading.

### Social and material deprivation indices

In addition to the SEFI-2, two additional ABSMs inspired by the work of Pampalon et al. were recently calculated at MCHP. These indices reflected the concept of deprivation as articulated by Townsend et al., who proposed multiple types of deprivation, including material deprivation and social deprivation. For this analysis, six census variables thought to primarily reflect either material deprivation or social deprivation were selected. To measure material deprivation, income (average household income), education (proportion of people without high school diplomas) and employment (unemployment rate) were included. To measure social deprivation, three family household measures were used: proportion of people separated, divorced, or widowed; proportion of people living alone; and proportion of people who had moved within the previous 5 years. For this analysis, unemployment rate and proportion of movers (in the previous 5 years) were substituted for Pampalon et al.’s inclusion of employment ratio (material deprivation) and proportion of single-parent families (social deprivation). In Manitoba, past analyses have revealed that the proportion of single-parent families is much more strongly related to material than to social deprivation variables. As expected, a factor analysis of the six variables included indicated that they separated into two distinct groups (see Table 2).

### Table 1. Factor Loadings for Socio-economic Factor Index 2 (SEFI-2), Manitoba, Using the 2006 Canadian Census

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average household income</td>
<td>-0.79</td>
</tr>
<tr>
<td>Unemployment*</td>
<td>0.67</td>
</tr>
<tr>
<td>No high school graduation†</td>
<td>0.74</td>
</tr>
<tr>
<td>Single-parent families‡</td>
<td>0.64</td>
</tr>
</tbody>
</table>

* Unemployment rate for labour force population aged 15 years and older.
† Proportion of population 15 years and older without high school graduation.
‡ Proportion of single-parent families.
Source: Ref. 6.

### Table 2. Factor Loadings for Social and Material Deprivation, Manitoba, Using the 2006 Canadian Census

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Social Deprivation</th>
<th>Material Deprivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average household income</td>
<td>-0.41</td>
<td>-0.75</td>
</tr>
<tr>
<td>Unemployment</td>
<td>0.05</td>
<td>0.61</td>
</tr>
<tr>
<td>No high school graduation</td>
<td>-0.15</td>
<td>0.89</td>
</tr>
<tr>
<td>Separated, divorced or widowed</td>
<td>0.82</td>
<td>0.15</td>
</tr>
<tr>
<td>Live alone</td>
<td>0.87</td>
<td>0.12</td>
</tr>
<tr>
<td>Moved in the previous 5 years</td>
<td>0.74</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

Source: Ref. 6.

### Population health measures

Four population health measures were calculated using data from the Population Health Research Data Repository housed at MCHP. This is a comprehensive collection of administrative, registry, survey and other databases primarily covering residents of Manitoba. The administrative health database holds records for virtually all contacts with the provincial health care system, the Manitoba Health Services Insurance Plan (including physicians, hospitals, personal care homes, home care, and pharmaceutical prescriptions), of all registered individuals. The four measures are as follows.

#### Premature Mortality Rate (PMR)

PMR is a standardized (age- and sex-adjusted) rate of “premature” death, or death before the age of 75 years. PMR is highly correlated with morbidity and with self-rated health. Populations with a high PMR require more health care services than other populations, and some researchers have suggested that this is the single best indicator of health status and need for health care.

#### Life Expectancy

Life expectancy is the average number of years an individual of a given age is expected to live if current age- and sex-specific mortality rates remain stable.

#### Potential Years of Life Lost (PYLL)

PYLL measures the number of years of life lost before a given age in the population (age 75 in the following analyses). As an alterna-
A relative measure of premature mortality, PYLL gives greater weight to deaths occurring at a younger age than to those at older ages.

**Self-rated Health**

The last health status indicator to be examined is a subjective measure of health status calculated from the Canadian Community Health Survey cycles 1.1, 2.1 and 3.1 carried out in 2001, 2003 and 2005.* Self-rated health was assessed on the survey with the following question: “In general, would you say your health is: Excellent? Very good? Good? Fair? Poor?”

Respondents were told that health referred not only to the absence of disease or physical injury or disability, but also to mental and social well-being. Unlike the previous measures of health status, self-rated health may be more accessible to some researchers as it does not depend on health records or contacts with the health system. The proportion of respondents who indicated that their health was either excellent or very good was calculated. Given the complexity of calculating area-level values for a question from a complex survey sample, a more complete description of the methods for calculating rates of self-rated health can be found in Metge et al.6

**Comparing health measures and ABSM**s

Average values on the three ABSMs previously described (SEFI-2, social deprivation, material deprivation) and simple household income were calculated for 25 urban areas in Winnipeg and 55 rural areas outside of Winnipeg. Values on the ABSMs were determined by simply calculating the DA population weighted average of the factor analysis scores in the area. The ABSMs were then compared with the sex- and age-adjusted rates for the health measures (PMR, PYLL and self-rated health). Relative risks on the health outcomes were estimated for each region using a Poisson or negative binomial regression model, which is the appropriate method for count data (e.g., the count of deaths for PMR, the count of years lost for PYLL). The log of the population in each stratum in the analysis was included in the model as an offset, so that the results accounted for the different population sizes in different areas. The rate of people reporting excellent or very good health was directly standardized to the weighted survey population for CCHS 2.1 and 3.1 cycles combined. All data management, programming and analyses were performed using SAS® software version 9.1 of the SAS System for Unix.

Figure 1 compares SEFI-2 with PMR for the City of Winnipeg. The 12 Winnipeg community areas are listed in order of increasing PMR, with the healthiest areas at the top and the least healthy areas at the bottom. There is remarkable consistency in the ordering of PMR and SEFI-2, the (mostly) suburban areas of the city having low ABSM scores and low PMRs. In contrast to the low scores for the suburbs, the core areas of Downtown and Point Douglas have higher PMR and considerably higher ABSM scores. This strong relation between the ABSM and health status is also present for the rural regional health authority areas (RHAs): the southern RHAs of Manitoba have better health and SES, whereas the more remote northern areas are worse off in both regards. The consistency of the relation in both urban and rural settings is an important finding, because urban DAs contain more relatively homogenous populations than rural areas.

A complete correlation matrix of the full set of ABSM and health status indicators is presented in Table 3. As with SEFI-2 and PMR, the correlations are quite high (e.g., r=0.76 for PYLL and material deprivation), indicating that most measures of ABSM and health are highly related, the exceptions being social deprivation and, to some degree, self-rated health. A factor analysis was also conduct-

---

ed on the full set of ABSM and health measures, and the results are presented in Table 4.

It might be expected that this factor analysis would result in high loadings for either the ABSM variables or the health status indicators, since these groups of variables would be expected to be highly correlated among themselves. The reality is that a variable from each set is represented in the two highest loadings (SEFI-2 and PMR). In addition, it would appear that any of several measures may be equally adequate at measuring the underlying concept of health status (PMR, life expectancy, PYLL) and that composite ABSMs (SEFI-2, material deprivation) can functionally be used as proxies of health status in relatively small populations. However, not all measures have such high loadings. The social deprivation index, in particular, has quite a low loading, and this is not surprising after checking the correlation matrix (Table 3). The reasons for these low correlations and low factor loading are difficult to discern, but according to Pampalon et al., it appeared that deprivation indices were not as strongly related to health in rural areas as in urban areas. Their results grouped material and social deprivation together, however. It may be the case that social deprivation indices, in particular, are a less accurate measure in rural areas. This is an issue that should be examined in closer detail in future research.

The factor that was identified from the analysis of the set of indicators can be interpreted as the underlying aspect of overall health (indicated by PMR, life expectancy, PYLL, self-rated health) that is associated with the indicators of SES, or social determinants of health (SEFI, material deprivation, social deprivation, income). Aside from the social deprivation index, the survey measure of health status (self-rated health) and the direct SES measure (average household income) have the lowest loadings in the factor analysis. While the composite ABSM and most health indicators share over 80% of their variance with the factor, household income shares only about 50% of the variance with the factor. This substantial difference suggests that average household income, though certainly not a poor measure of SES, is not as statistically efficient a measure as SEFI-2 or material deprivation is, and this may be particularly true when relating SES to health.

**CONCLUSIONS**

ABSMs are a remarkably efficient means of estimating SES for small areas, and the availability of information from the Canadian cen-

### Table 3. Correlations Between Health and SES Sentinel Indicators: Manitoba

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Spearman Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PYLL Life Expectancy‡ SEFI-2 Life expectancy‡ Excellent/very good self-rated health Material deprivation Social deprivation Average household income</td>
</tr>
<tr>
<td>PMR</td>
<td>0.86** -0.93**</td>
</tr>
<tr>
<td>PYLL</td>
<td>-0.86** -0.93**</td>
</tr>
<tr>
<td>Life expectancy‡</td>
<td>0.51** 0.55**</td>
</tr>
<tr>
<td>Excellent/very good self-rated health</td>
<td>0.80**</td>
</tr>
<tr>
<td>SEFI-2</td>
<td>0.83** 0.85**</td>
</tr>
<tr>
<td>Material deprivation</td>
<td>0.57** 0.59**</td>
</tr>
<tr>
<td>Social deprivation</td>
<td>0.87** 0.89**</td>
</tr>
<tr>
<td>Average household income</td>
<td>0.86** 0.88**</td>
</tr>
</tbody>
</table>

**p < 0.05  * p < 0.01  ** p < 0.001
‡ Only life expectancy (calculated at birth) for males was included in this table; correlation patterns were similar for females.
PYLL = potential years of life lost; PMR = premature mortality rate.
Source: Ref. 6.

### Table 4. Factor Loadings for the Health Status Composite Index

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Factor 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMR</td>
<td>0.93</td>
</tr>
<tr>
<td>SEFI-2</td>
<td>0.97</td>
</tr>
<tr>
<td>PYLL</td>
<td>0.91</td>
</tr>
<tr>
<td>Life expectancy‡</td>
<td>-0.90</td>
</tr>
<tr>
<td>Social deprivation</td>
<td>-0.26</td>
</tr>
<tr>
<td>Material deprivation</td>
<td>0.92</td>
</tr>
<tr>
<td>Self-rated health (excellent/very good)</td>
<td>-0.73</td>
</tr>
<tr>
<td>Average household income</td>
<td>-0.72</td>
</tr>
</tbody>
</table>

PMR = premature mortality rate; PYLL = potential years of life lost.
Source: Ref. 6.

sus makes this possible not only in Manitoba but also across the country. Composite indices such as SEFI-2 and material deprivation appear to be more robust measures of SES than income alone, at least so far as this concept relates to the social determinants of health. Where income may vary for idiosyncratic reasons (e.g., in northern mining communities), the inclusion of education in the component may temper the impact of extreme values to get a better representation of SES as it relates to the social determinants of health. Further highlighting the robustness of these measures is that modifications of the variables included (e.g., material deprivation here as compared with the variables of Pampalon et al.) appear to have little impact on the strength of the ABSM-health status relation. On the other hand, on the basis of the correlation matrix presented in Table 3, if one is simply intent on adjusting for SES, an area-based income measure may suffice and is much more straightforward to include. The use of income alone also avoids any issues that may be related to the coherence of a composite ABSM; in Manitoba, for instance, employment ratio and proportion of single-parent families did not show patterns similar to those of earlier work in Quebec, and alternative indicators had to be substituted.

Although work at MCHP with ABSMs has primarily been conducted on populations of people, evaluations of SEFI and similar ABSMs suggest that information at the DA level is fine grained enough to reduce concerns about the modifiable areal unit problem and that ABSMs independently contribute to health beyond individual-level assessments of SES. Including these types of measures in analyses of health outcomes can help planners properly evaluate the potential impact of policies that do, or do not, address the impact of SES on health.
REFERENCES


