ON-Marg
Ontario Marginalization Index

User Guide
version 1.0
Centre for Research on Inner City Health
St. Michael’s Hospital
209 Victoria Street
Toronto, ON M5B 1W8

Developers
Flora I. Matheson, PhD
James R. Dunn, PhD
Katherine L. W. Smith, MHSc
Rahim Moineddin, PhD
Richard H. Glazier, MD, MPH

Contact
mathesonf@smh.ca
The Ontario Marginalization Index (ON-Marg) is a census- and geographically-based index that seeks to:

- show differences in marginalization between areas; and
- understand inequalities in various measures of health and social well-being, either between population groups or between geographical areas.

ON-Marg is an Ontario-specific version of the Canadian Marginalization Index (CAN-Marg), which has been in use since 2006 and is available from www.canmarg.ca.

ON-Marg is multifaceted, allowing researchers and policy and program analysts to explore multiple dimensions of marginalization in urban and rural Ontario.

The four dimensions are:

- residential instability
- material deprivation
- ethnic concentration
- dependency

The index was developed using a theoretical framework based on previous work on deprivation and marginalization. It was then empirically derived using principal components factor analysis. It has been demonstrated to be stable across time periods and across different geographic areas (e.g. cities and rural areas). It has also been demonstrated to be associated with health outcomes including hypertension, depression, youth smoking, alcohol consumption, injuries, body mass index and infant birthweight.1-7

ON-Marg can be used for:

1) Planning and needs assessment. For example, if the goal is to identify service gaps, ON-Marg can be used to identify where rates of hospitalizations for a particular disease, such as diabetes, are high and additional services might be needed.

2) Resource allocation. For example, marginalization indexes could be used in funding formulae for primary health care services.

3) Monitoring of inequities. For example, marginalization indexes can provide a way to monitor changes in areas over time to look for improvement or to identify areas that may be in decline.

4) Research. For example, in the health sector there is a long history of using small area indexes to describe the relationship between marginalization and health outcomes; greater marginalization is associated with higher mortality rates, and higher rates of many diseases.8-12
Following a literature review, 42 variables were selected from the 2001 Canadian census for potential inclusion in the index (see Appendix I). Principal component factor analysis yielded four factors with Eigenvalues greater than 1. Of the original variables, 18 were included in the four factors (see Table 1). The 2001 index was created from two core files with 49,153 dissemination areas (DAs) and 4,757 census tracts (CTs). The index was replicated using 2006 data with 52,973 DAs and 5,017 CTs.

Factor loadings were used to compute a separate index for each of the four dimensions. Each dimension is an asymmetrically standardized scale.

ON-Marg applies to areas, not individual people. Scores for each dimension are available for every census tract and dissemination area in Ontario, except where data is suppressed.

ON-Marg is available for download in Excel 2002 format for the 2001 and 2006 census years. Files are named:

- ON-Marg_2001.xls
- ON-Marg_2006.xls
- ON-Marg_geographic_crosswalk_2001.xls
- ON-Marg_geographic_crosswalk_2006.xls

The 2001 and 2006 ON-Marg files have tabs for the following geographies:

- dissemination areas
- census tracts
- census divisions
- census subdivisions
- LHINs
- sub-LHINs
- public health units
- consolidated municipal service manager areas

ON-Marg values for geographies larger than census tracts were derived from the original DA factor scores.

Quintile values are not provided for LHINs, public health units, census divisions and consolidated municipal service manager areas because there are too few geographic units within each area, which disguises heterogeneity. For example, a LHIN with 14 CTs or DAs has too few geographic units to create meaningful quintiles.

A dissemination area (DA) is a small, relatively stable geographic unit composed of one or more adjacent dissemination blocks. It is the smallest standard geographic area for which all census data are disseminated. DAs cover all the territory of Canada. 

A census tract (CT) is a small, relatively stable geographic unit with a population of 2,500 to 8,000 people constructed similarly with respect to economic status and social conditions. Census tracts are located in census metropolitan areas and in census agglomerations having an urban core population of 50,000 or more as of the most recent census.
### Table 1. Dimensions of marginalization and their respective indicators

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential instability</td>
<td>Proportion of the population living alone</td>
</tr>
<tr>
<td></td>
<td>Proportion of the population who are not youth (aged 16+)***</td>
</tr>
<tr>
<td></td>
<td>Average number of persons per dwelling***</td>
</tr>
<tr>
<td>Material deprivation</td>
<td>Proportion of the population aged 20+ without a high-school diploma**</td>
</tr>
<tr>
<td></td>
<td>Proportion of families who are lone parent families</td>
</tr>
<tr>
<td></td>
<td>Proportion of the population receiving government transfer payments</td>
</tr>
<tr>
<td></td>
<td>Proportion of dwellings that are apartment buildings</td>
</tr>
<tr>
<td></td>
<td>Proportion of the population who are recent immigrants (arrived in the 5 years prior to census)</td>
</tr>
<tr>
<td>Dependency</td>
<td>Proportion of the population who are aged 65 and older</td>
</tr>
<tr>
<td></td>
<td>Dependency ratio (total population 0-14 and 65+ /total population 15 to 64 )</td>
</tr>
<tr>
<td></td>
<td>Proportion of the population not participating in labour force (aged 15+)***</td>
</tr>
<tr>
<td>Ethnic concentration*</td>
<td>Proportion of the population who self-identify as a visible minority</td>
</tr>
</tbody>
</table>

* Aboriginal indicators did not load on any of the factors.

** For the 2006 index, the indicator is the proportion of the population aged 25+ without a certificate, diploma or degree. This is due to a change in the Statistics Canada definition.

*** Indicators were reverse coded, meaning they were coded opposite of the measure (e.g. % married/common law becomes %single/divorced/separated/widowed).

**** “Low income” is defined as below the low income cutoff (LICO), a Statistics Canada measure that is adjusted for community size, family size and inflation.
The ON-Marg dimensions can be used separately or combined into a composite index (see the next section). Whether you use individual dimensions or the combined index will be determined by the research question.

For each dimension, ON-Marg is provided in two forms:

I. Factor scores (interval scale): Factor scores are constructed from the principal component factor analysis. They were derived from the Canadian Marginalization Index (CAN-Marg, see www.canmarg.ca) and, when the full Canadian index is used, represent a standardized scale with a mean of 0 and a standard deviation of 1. Lower scores on each dimension correspond to areas that are the least marginalized; higher scores on each dimension correspond to areas that are the most marginalized.

II. Quintiles (ordinal scale): Quintiles have been created by sorting the marginalization data into five groups, ranked from 1 (least marginalized) to 5 (most marginalized). Each group contains a fifth of the geographic units. For example, if an area has a value of 5 on the material deprivation scale, it means it is in the most deprived 20 percent of areas in Ontario.

The objectives of your analysis and the methods you are using will determine whether you use factor scores or quintiles in your analysis. For example, a mapping exercise might be best presented using quintiles, whereas a regression model might benefit from the detail of the factor scores.

**SUMMARY SCORE FOR THE ON-MARG DIMENSIONS**

Users may wish to examine overall marginalization using a summated score. This can be done using the quintile scales for each dimension.

**In order to calculate the summated score, follow these steps:**

1. Compare the correlations between each dimension with the outcome. This allows you to determine if the associations are in the same direction (either all positive or all negative). If the associations are either all positively or all negatively associated with the outcome then an average marginalization score can be computed. If one or more dimensions are in the opposite direction it is not recommended to combine the dimensions. For example, if ethnic concentration is negatively associated with the outcome of interest, this may represent a protective factor (e.g. a healthy immigrant effect) and it may not be appropriate to combine ethnic concentration with the other dimensions that are positively associated with the outcome and therefore represent risk factors.
Sum the quintile values across the four dimensions.

Divide by 4 (which is the number of dimensions).

These steps will produce a score ranging from 1 to 5 where 1 reflects low levels of marginalization and 5 reflects high levels of marginalization.

\[
\text{Summary Score} = \frac{\text{instability\_quintile} + \text{deprivation\_quintile} + \text{dependency\_quintile} + \text{ethnic\_con\_quintile}}{4}
\]

**Caution:** Factor scores cannot be used to obtain a summary score.

**CALCULATING AN AVERAGE ON-MARG SCORE VALUE FOR HIGHER-ORDER GEOGRAPHICAL UNITS**

Commonly-used Ontario geographies have already been aggregated and are available as part of the ON-Marg package.

However, some research and policy questions require geo-coding at custom geographic units. You can use the DA and CT data in ON-Marg and the methods described in this section to create values for your own geographies, using population-weighted average scores.

**Example: calculating weighted average scores for a single Ontario urban health region from 2006 CT- or DA-level marginalization scores.**

1. Define the health region in terms of the component CTs and/or DAs.

2. Using the population counts, take the weighted average of each factor score value across all the CTs or DAs in the health region. Use the CT (or DA, depending on your analysis) worksheet of the ON-Marg file ON-Marg_2006.xls to obtain the weighted average for the health region, following these steps:
   a. Multiply each CT or DA marginalization score value by the population within the CT or DA for the health region.
   b. Sum the multiplied values from a). This becomes the numerator.
   c. Sum the population values from each CT or DA to obtain a total population count for the health region. This becomes the denominator.
   d. Divide the total from (b) by the total from (c). This is your weighted average.

Weighted average deprivation score:

\[
\frac{\sum(\text{ON-Marg\_CT\_2006} \times \text{ONPop\_CT\_2006})}{\sum(\text{ONPop\_CT\_2006})}
\]

3. You can now use these weighted averages to create quintiles.

**Caution:** Weighted averages can disguise heterogeneity within large geographic areas. For example, when the weighted average method is used to determine the deprivation quintile for the South-east sub-LHIN in Toronto, the result is 5 (most deprived). Figure 1, however, shows the true variation in this sub-LHIN by using summed DA population counts by quintile, not weighted averages, to show the number of people in each quintile. The resulting graph shows there are pockets of low, moderate and high deprivation in the South-east sub-LHIN that would be masked by using the summary score of 5.
Figure 1. Population in each quintile in South-east sub-LHIN of the Toronto Central LHIN, based on DA population

Using on-marg for analysis

I. EXPLORING THE RELATIONSHIP BETWEEN OUTCOMES AND AREA-LEVEL MARGINALIZATION

Outcomes can include the following:

• individual health status;
• individual risk or protective factors;
• rates of disease, or any health related event.

Research questions that could be answered include:

1) What is the association between health outcomes, such as mortality and diabetes rates, and area-level marginalization?

2) What is the association between health behaviours, such as smoking and alcohol consumption, and area-level marginalization?

3) What is the association between access to routine surgical procedures, such as joint replacement, and area-level marginalization?

To answer such questions, merge the outcome file with ON-Marg, following the steps below:

1) Prepare the outcome file:
   a) Ensure the addresses are error-free.
   b) Geocode each observation in your outcome data set (e.g. mortality, crime events, hypertension) to CT or DA. Often this is accomplished using the PCCF4+ SAS program created by Statistics Canada. Now every record is associated with a particular CT or DA.

* Health behaviours such as smoking and drinking are often examined as risk and/or protective factors, but can also be outcomes of interest.
2. Merge your health outcome data set with the ON-Marg CT or DA, thus linking each geocoded outcome with the appropriate area marginalization scores.

II. USING ON-MARG AS AN INDIVIDUAL-LEVEL PROXY

In some instances, ON-Marg can be used as a proxy for individual-level data when actual data is not available. If individual-level socio-economic status data is unavailable, for example, DA-level factor scores or quintiles for deprivation can be assigned to each individual based on the DA in which the individual resides and used as a proxy for socioeconomic status.

Caution: To minimize measurement error, use the smallest spatial area available. In the case of ON-Marg, this is DA data. The reason is similar to that provided under the “caution” for weighted averages on page 6. As the size of the geographic unit increases (e.g. CTs and sub-LHINs), the potential for ecological fallacy increases as well, since not everyone in a marginalized area is marginalized.

In effect, using areas larger than the DA will weaken any relationship between individual- and area-level marginalization. The larger the geographic area, the less likely it is that an individual’s socio-economic status will actually correspond to the deprivation score of the area in which s/he lives.

III. MAPPING THE INDEX

The index can be displayed geographically using mapping software such as ArcGIS or MapInfo.

IV. COMPARING THE MARGINALIZATION OF 2 OR MORE GROUPS

If you want to compare levels of marginalization between two or more groups (e.g. hypertensive versus non-hypertensive; diabetic versus non-diabetic) you can compare the distributions of quintiles (or factor scores) using a non-parametric test. This test is used because quintile values are ordinal, and the principal component scores are skewed.

V. COMPARING RATES OF EVENTS

If you are comparing rates of events with marginalization (e.g. mortality rates in a region compared across the five marginalization scale values) you can calculate a rank correlation coefficient, or simply plot your results. Note that the denominators for your rates can be obtained from the CT or DA populations.

ethics & confidentiality

Ethical approval for the development of ONMARG 2001 and 2006 was obtained in April 2005 from the Research Ethics Board of St. Michael’s Hospital.

Access to the data used in this study was provided by Statistics Canada under the Data Liberation Initiative at the University of Toronto.
Missing data: There is some missing data in the DA and CT files due to data suppression (e.g. income). Additionally, in some areas, input variables have a value of 0. For example, a DA may not have any recent immigrants.

Time period of data: Data for the index is from the 2001 and 2006 census years and users should be aware of this when selecting the most appropriate year for their own analyses. For example, if your outcome data was collected in 2005 or 2007 you would use the 2006 index to ensure data comparability. If your outcome data was collected in 2000 or 2002, you would use the 2001 index.

Coverage of census: Some populations, for example Aboriginal people living on reserves, may be under-counted in the census. ON-Marg may not be as sensitive for these populations. Refer to Statistics Canada to see if census coverage will impact your analyses.

Acknowledgements

We gratefully acknowledge the contributions of our advisory panel members who helped us develop this user guide. We would also like to acknowledge and thank Public Health Ontario (formerly the Ontario Agency for Health Protection and Promotion) and the Canadian Population Health Initiative (Canadian Institute for Health Information).

Advisory Panel

Chris Borst
City of Hamilton, Social Development & Early Childhood Services
Lisa Corscadden
Canadian Population Health Initiative
Cynthia Damba
Toronto Central Local Health Integration Network
Etan Diamond
Ontario Municipal Social Services Association
Eric Holowaty
Population Studies and Surveillance, Cancer Care Ontario
Marc Lefebvre
Centre de Santé Communautaire du Grand Sudbury
Sara Mayo
Hamilton Social Planning and Research Council

Caitlin Mills
Ontario Health Study
Todd Norwood
Cancer Prevention and Control, Cancer Care Ontario
Jane Polsky
Centre for Research on Inner City Health
Paul Ritvo
Cancer Prevention and Control, Cancer Care Ontario
Ruth Sanderson
Public Health Ontario
Michael Spinks
South East Local Health Integration Network
Julie Stratton
Region of Peel, Public Health
Susitha Wanigaratne
Cancer Prevention and Control, Cancer Care Ontario
### Appendix I: Census Variables

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Proportion of the population who moved during the past 5 years</td>
</tr>
<tr>
<td>2</td>
<td>Proportion living in same house as 1 year ago</td>
</tr>
<tr>
<td>3</td>
<td>Proportion of population lone parent families</td>
</tr>
<tr>
<td>4</td>
<td>Proportion of population living alone</td>
</tr>
<tr>
<td>5</td>
<td>Dependency ratio (total population 15 to 64/total population 0-14 and 65+)</td>
</tr>
<tr>
<td>6</td>
<td>Proportion of population youth (aged 5-15)</td>
</tr>
<tr>
<td>7</td>
<td>Proportion foreign born</td>
</tr>
<tr>
<td>8</td>
<td>Proportion Aboriginal</td>
</tr>
<tr>
<td>9</td>
<td>Proportion of the population who are recent immigrants (arrived in the 5 years prior to census)</td>
</tr>
<tr>
<td>10</td>
<td>Proportion with no official language</td>
</tr>
<tr>
<td>11</td>
<td>Proportion unemployed (aged 15+)</td>
</tr>
<tr>
<td>12</td>
<td>Labour force participation rate (aged 15+)</td>
</tr>
<tr>
<td>13</td>
<td>Proportion who self-identify as a visible minority</td>
</tr>
<tr>
<td>14</td>
<td>Proportion aged 15-24 not attending school</td>
</tr>
<tr>
<td>15</td>
<td>Proportion aged 20+ without high school diploma</td>
</tr>
<tr>
<td>16</td>
<td>Proportion of the population considered low income using the low income cutoff (LICO)</td>
</tr>
<tr>
<td>17</td>
<td>Average household income</td>
</tr>
<tr>
<td>18</td>
<td>Proportion of the population receiving government transfer payments</td>
</tr>
<tr>
<td>19</td>
<td>Proportion with no religious affiliation</td>
</tr>
<tr>
<td>20</td>
<td>Average dollar value of dwelling</td>
</tr>
<tr>
<td>21</td>
<td>Proportion of dwellings that are apartment buildings</td>
</tr>
<tr>
<td>22</td>
<td>Proportion of owner households spending 30% or more of household income on major payments</td>
</tr>
<tr>
<td>23</td>
<td>Proportion of tenant households spending 30% or more of household income on rent</td>
</tr>
<tr>
<td>24</td>
<td>Proportion of dwellings that are owned</td>
</tr>
<tr>
<td>25</td>
<td>Proportion of occupied units that are rentals</td>
</tr>
<tr>
<td>26</td>
<td>Proportion of population self-employed</td>
</tr>
<tr>
<td>27</td>
<td>Proportion of population female</td>
</tr>
<tr>
<td>28</td>
<td>Proportion of population married/common law</td>
</tr>
<tr>
<td>29</td>
<td>Proportion of households living in dwellings that are in need of major repair</td>
</tr>
<tr>
<td>30</td>
<td>Proportion of population aged 15+ doing unpaid housework</td>
</tr>
<tr>
<td>31</td>
<td>Proportion of population aged 15+ looking after children without pay</td>
</tr>
<tr>
<td>32</td>
<td>Proportion of population aged 15+ providing unpaid care/assistance to seniors</td>
</tr>
<tr>
<td>33</td>
<td>Raw population count</td>
</tr>
<tr>
<td>34</td>
<td>Average number of persons per dwelling</td>
</tr>
<tr>
<td>35</td>
<td>Average number of persons per room</td>
</tr>
<tr>
<td>36</td>
<td>Ratio of employment to population</td>
</tr>
<tr>
<td>37</td>
<td>Average income</td>
</tr>
<tr>
<td>38</td>
<td>Proportion of persons separated, divorced or widowed</td>
</tr>
<tr>
<td>39</td>
<td>Proportion of children younger than 6 years</td>
</tr>
<tr>
<td>40</td>
<td>Persons per square kilometer</td>
</tr>
<tr>
<td>41</td>
<td>Unemployment rate in private households with children under 6 years</td>
</tr>
<tr>
<td>42</td>
<td>Proportion of the population who are aged 65 and older</td>
</tr>
</tbody>
</table>

Shading represents variables chosen for use in ON-Marg.


