Immigration and all-cause mortality in Canada: An illustration using linked census and administrative data

Seminar presentation, Quebec Interuniversity Centre for Social Statistics (QICSS), November 26, 2014

Walter Omariba, PhD
Acknowledgments

• Edward Ng, PhD, Health Analysis Division, Statistics Canada
• Bilkis Vissandjée, School of Nursing, University of Montreal
• Michael Tjekema, Health Analysis Division, Statistics Canada.
Outline

- Background
  - Immigration trends in Canada
  - Explaining immigrant mortality advantage
  - Limitations of previous mortality studies

- Study’s questions

- Record linkage and benefits

- 1991 census cohort description
  - Linkage results
  - Potential research areas

- Study sample and analytical methods

- Results

- Limitations and strengths

- Conclusion
Background

- Even though Canada is historically an immigrant country, immigration is increasingly playing an important role in the country’s demographic profile.

- In the 2006 Census 19.6% of the population was foreign-born and increased to 20.6 in the 2011 NHS.
  - Projected to reach between 25% and 28% by 2031 (Malenfant et al. 2009).
  - Between 2001 and 2006, newcomers comprised 69.3% of the people added to the population; this had declined slightly to 62.4% between 2006 and 2011.

- There is also a shift in the source countries from Europe to mostly Asia.
Background continued

Table 1: Top five birthplace of recent immigrants, 1981 to 2011

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UK</td>
<td>Hong Kong</td>
<td>China</td>
<td>China</td>
<td>Philippines</td>
</tr>
<tr>
<td>2</td>
<td>Vietnam</td>
<td>Poland</td>
<td>India</td>
<td>India</td>
<td>China</td>
</tr>
<tr>
<td>3</td>
<td>USA</td>
<td>China</td>
<td>Philippines</td>
<td>Philippines</td>
<td>India</td>
</tr>
<tr>
<td>4</td>
<td>India</td>
<td>India</td>
<td>Pakistan</td>
<td>Pakistan</td>
<td>USA</td>
</tr>
<tr>
<td>5</td>
<td>Philippines</td>
<td>Philippines</td>
<td>Hong Kong</td>
<td>USA</td>
<td>Pakistan</td>
</tr>
</tbody>
</table>

**Note:** 'Recent immigrants' refers to landed immigrants who arrived in Canada within five years prior to a given census.

**Sources:** Statistics Canada, censuses of population, 1981 to 2011
Brown – Asian Country
Green – Europe or United States
Background continued

- Given the changing demographic profile of Canada, it is critical to understand the health risks associated with immigration as well as healthcare utilisation.

- Overall, immigrants tend to have better health outcomes (mortality, morbidity, hospitalisation) compared to non-immigrants.

- Based on review of literature, there are several explanations for the immigrant mortality advantage:
  - Healthy immigrant effect,
  - Data artefact, and
  - Cultural effects.
Explaining immigrant mortality advantage

- Healthy immigrant effect: Immigrants are selected for better health at the outset: Health enhancing characteristics and/or better physical and mental health (e.g., Hajat et al. 2010).

- Data artefact: data quality (e.g., Palloni & Arias 2004) and the ‘salmon bias’ (Pablos-Mendez 1994).

- Cultural effects: Health behaviours and interaction with the environment (Franzini et al. 2001; Abraido-Lanza et al. 2005; Viruell-Fuentes & Schulz 2009).
Limitations of previous mortality studies

- The testing of these hypotheses is hampered by lack of data:
  - Administrative data: Details about deaths, age and sex.
  - Census or survey data: Characteristics of individuals including immigrant status, but no information on deaths.
- Concurrent examination of country of birth, period of immigration and relevant predictors was not possible in previous studies.
- Linked data such as the 1991 Canadian Census Cohort Mortality & Cancer Follow-up Study address these limitations.
Research questions and goal

- Q1. Do immigrants have a mortality advantage compared to the Canadian-born?
- Q2. If immigrants have a mortality advantage, does it decline as their duration of residence in Canada increases? Is this dependent on age?
- Q3. What is the role of socioeconomic and sociodemographic factors on the observed immigrant mortality patterns?

Goal: Highlight the availability and utility of the 1991 Canadian Census Mortality and Cancer Follow-up Study.
Why data linkage?

- Administrative data in Canada do not uniformly contain individual identifiers (socioeconomic status, ethnicity, Aboriginal) or other characteristics beyond basic demographic information (age, sex, residence).

- Few datasets are suitable for geographic linkage with environmental exposure data due to lack of detailed place of residence information.

- Difficult to provide health indicators for important population sub-groups.
What is record linkage?

- Combines two or more datasets using common identifiers
  - Deterministic
  - Probabilistic.

- Need to achieve a balance between the need to protect privacy of individuals and the public good a linkage may achieve.
Benefits of a census linkage

- **Expanded knowledge base**
  - Improved understanding of social determinants.
  - Allow for multi-variable & multi-level analysis.
  - Environmental exposure studies.
  - Identification of multiple dimensions of socioeconomic disadvantage
    - With respect to education, income, occupation, housing, etc…

- **Large cohort size**
  - Analysis of population sub-groups
    - Such as immigrants, marginally housed, ethnic origins, First Nations, Métis, and Inuit.
  - Ability to examine rare outcomes.
  - Allow for cross-classification
    - Urban – Aboriginal; Cardiovascular Disease – Recent immigrants.
1991 census cohort

- **Purpose of the linkage:** Develop a set of baseline indicators of mortality to monitor health inequalities.

- **Eligibility:**
  - Enumerated on 1991 census long form (1 in 5 (20%) households *).
  - Aged 25 or older as of June 4, 1991.
  - Not a usual resident of an institution.
  - Linkage approval for 15% of persons aged 25+.

- Note that 3.4% of the Canadian population of all ages were not enumerated by the 1991 census.

* Note that all residents of Indian Reserves and remote northern communities receive long form questionnaire.
Structure of the 1991 Canadian Census Cohort

- Canadian Mortality Database: 1991-2011
- Tax-filer data: 1984-2011
- Canadian Cancer database (CCDB): 1969-2011
- Longitudinal Worker File: 1983-2011

Linkage approval to 2011

1991 Census Cohort

Source: Peters et al. 2013
Content

- 1991 Census
  - Demography, labour market, income, education, language, disabilities, housing, immigration, ethno-cultural, Aboriginal ancestry, Registered Indian.

- Tax-filer Summary File (T1 Family File (T1FF))
  - Annual place of residence (postal code on tax return), marital status- tracking of mobility.

- Canadian Cancer database (CCDB):
  - Diagnosis site of primary malignant neoplasm, morphology, topology, date and province of diagnosis, date of death.

- Canadian Mortality Database (CMDB)
  - Underlying cause of death, date of death, age at death.

- Longitudinal Worker File (LWF).
  - Employment income, history, and reason of job separation.
1991 census cohort

Cohort creation

- Eligible census respondents linked to tax filer data (non-financial) in order to get names.
- Matching variables: sex, date of birth, postal code, spousal date of birth.
- Results: 80% linkage rate, 99% correct links.

Deterministic linkage of LWF to tax summary file for annual place of residence.

- Postal codes (1984-2008), approval to 2011.
- Employment history (1983-2010), approval to 2011.

Probabilistic linkage to mortality and cancer.

- Matching variables: sex, date of birth, names, postal code.
- Cancer (1969 to 2003), approval to 2011.
Table 2: In-scope* and cohort breakdown

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>In-scope</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (count)</td>
<td>3,576,485</td>
<td>2,734,835</td>
</tr>
<tr>
<td><strong>Sex (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>48.6</td>
<td>49.7</td>
</tr>
<tr>
<td>Female</td>
<td>51.4</td>
<td>50.3</td>
</tr>
<tr>
<td><strong>Age (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 to 44</td>
<td>52.6</td>
<td>54.5</td>
</tr>
<tr>
<td>45 to 64</td>
<td>30.5</td>
<td>30.0</td>
</tr>
<tr>
<td>65 +</td>
<td>16.9</td>
<td>15.4</td>
</tr>
<tr>
<td><strong>Educational attainment (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than secondary graduation</td>
<td>37.8</td>
<td>34.9</td>
</tr>
<tr>
<td>Secondary graduation or higher</td>
<td>62.2</td>
<td>65.1</td>
</tr>
<tr>
<td><strong>Income adequacy quintile (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quintile 1-poorest</td>
<td>20.0</td>
<td>17.2</td>
</tr>
<tr>
<td>Quintile 5-richest</td>
<td>20.0</td>
<td>21.5</td>
</tr>
</tbody>
</table>

* In-scope refers to all individuals who were enumerated by the long-form, were aged 25+, and were not a resident of an institution

Source: Peters et al. 2013
Linkage results

1991 Census Cohort (n=2,734,835)

Longitudinal Worker File: 1983-2010 (n=264,010)

Cancer: 1969-2003 (cases=338,085)

Mobility: 1984-2006 (followed=2,643,769)

Mortality: 1991-2006 (deaths=426,979)

Linkage approval to 2011

Source: Peters et al. 2013
Results - survival

Figure 1: Percentage surviving to various ages in Canada for 1995-1997 and 2002 (average) compared to cohort for 1991-2006

Source: Peters et al., 2013
Results - Cancer incidence

Figure 2: Age-standardized incidence rates of cancer, the cohort compared to Canadian Cancer Registry

Source: Peters et al., 2013
Potential research areas

- **Sub-population analysis**
  - First Nations, Métis, immigrants (year of immigration), place of birth, ethnic origin etc.

- **Analysis by socioeconomic status**
  - Income (source, household, individual), education (years, qualifications), occupation, industry, type of housing, marital status.

- **Multi-dimensional analysis**

- **Exposure analysis**
  - Assign exposure via postal code representative points.

- **Labour outcomes**
  - Economic outcomes associated with cancer survival.
Study sample

- The 1991 CCMCFS:
  - The first follow-up: 1991-2001 (No cancer data)
  - Follow-up period for the study: 1991-2006 (Cancer and employment data)

- Sample description:
  - Cohort sample: N=2,734,835.
  - Analysis sample: n=2,719,500.
    - Exclusions: non-permanent residents (n=14,300) and people born in Canada classified as immigrants (n=1,000).
Variables

- Outcome variable: Risk of death measured by duration of survival in the follow-up period.
  - Deaths included in the analysis: 425,785.

- Independent variables: Immigrant status and duration in Canada.
  - Control variables: age, marital status, knowledge of official languages, education, income quintiles, and employment.
Analytical methods

- Cox proportional hazard model used:
  - Conditional on survival to time $t$, the model estimates a non-parametric baseline risk of death at time $t$ for individual $i$.
  - The focus is mainly on the predictors and less on shape of the baseline hazard.

- Models were estimated separately for males and females and selected countries (UK, India, China/Hong Kong, Philippines, and the Caribbean)

- We examined separately, differences by immigration status and duration of residence.
### Table 3: Description of the sample

<table>
<thead>
<tr>
<th></th>
<th>Non-immigrants</th>
<th>Total immigrants</th>
<th>Immigrants</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UK</td>
<td>China/HK</td>
<td>India</td>
<td>Philippines</td>
</tr>
<tr>
<td>Both sexes</td>
<td>2,167,200 (79.9%</td>
<td>552,300 (20.3%)</td>
<td>100,700</td>
<td>37,000 (18.6%)</td>
<td>21,100 (10.6%)</td>
<td>14,800 (7.4%)</td>
<td>25,100 (12.6%)</td>
</tr>
<tr>
<td>Number (%)</td>
<td>335,000</td>
<td>90,800</td>
<td>25,200</td>
<td>3,500 (10.7%)</td>
<td>1,600 (4.7%)</td>
<td>1000 (2.9%)</td>
<td>1,900</td>
</tr>
<tr>
<td>All cause deaths (%)</td>
<td>(78.7%)</td>
<td>(21.3%)</td>
<td>(75.9%)</td>
<td>(10.7%)</td>
<td>(4.7%)</td>
<td>(2.9%)</td>
<td>(5.7%)</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-44</td>
<td>57.0</td>
<td>44.1</td>
<td>32.0</td>
<td>58.1</td>
<td>59.1</td>
<td>63.8</td>
<td>56.8</td>
</tr>
<tr>
<td>45-64</td>
<td>28.3</td>
<td>37.1</td>
<td>38.2</td>
<td>30.0</td>
<td>33.5</td>
<td>28.4</td>
<td>36.6</td>
</tr>
<tr>
<td>?65</td>
<td>14.7</td>
<td>18.8</td>
<td>29.8</td>
<td>11.9</td>
<td>7.4</td>
<td>7.7</td>
<td>6.6</td>
</tr>
<tr>
<td>Duration in Canada, %</td>
<td>...</td>
<td>18.6</td>
<td>5.7</td>
<td>43.9</td>
<td>30.5</td>
<td>41.5</td>
<td>20.7</td>
</tr>
<tr>
<td>&lt;10 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-19 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-34 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;=35 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source:** 1991-2006 Canadian Census Mortality and Cancer Follow-up Study
Figure 3: Age Standardised Mortality Rate (per 100,000 person years lived)

Source: The 1991 Canadian Census Cohort Mortality & Cancer Follow-up Study
Do immigrants have a mortality advantage?

Figure 4: Hazard ratios of mortality by sex, overall cohort, and selected countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>China/Hongkong</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philipinnes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caribbean</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ref: Non-immigrants

Note: All the ratios are statistically significant

Source: The 1991 Canadian Census Cohort Mortality and Cancer Follow-up Study
Table 4: Hazard ratios for all-cause mortality by immigrants duration in Canada compared to non-immigrants, 1991-2006 follow-up

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>95% CI</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10 years</td>
<td>0.60</td>
<td>0.58 - 0.62</td>
</tr>
<tr>
<td>10-19 years</td>
<td>0.67</td>
<td>0.65 - 0.69</td>
</tr>
<tr>
<td>20-34 years</td>
<td>0.75</td>
<td>0.74 - 0.77</td>
</tr>
<tr>
<td>&gt;=35 years</td>
<td>0.85</td>
<td>0.84 - 0.86</td>
</tr>
<tr>
<td>UK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 years</td>
<td>0.72</td>
<td>0.68 - 0.77</td>
</tr>
<tr>
<td>&gt;=20 years</td>
<td>0.87</td>
<td>0.86 - 0.89</td>
</tr>
<tr>
<td>China/Hong Kong</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 years</td>
<td>0.59</td>
<td>0.56 - 0.62</td>
</tr>
<tr>
<td>&gt;=20 years</td>
<td>0.66</td>
<td>0.61 - 0.71</td>
</tr>
<tr>
<td>Country</td>
<td>Age Group</td>
<td>Male Hazard Ratio</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>India</td>
<td>&lt;20 years</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>&gt;=20 years</td>
<td>0.60</td>
</tr>
<tr>
<td>Philippines</td>
<td>&lt;20 years</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>&gt;=20 years</td>
<td>0.60</td>
</tr>
<tr>
<td>Caribbean</td>
<td>&lt;20 years</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>&gt;=20 years</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Source: Same as Table 3
Is the duration effect dependent on age?

Figure 5: Hazard ratios of mortality by age and duration in Canada, all cohort

Source: Same as Table 3
Limitations

- Census characteristics measured at baseline.
- No lifestyle and proximate factors in the data such as smoking, alcohol drinking, engagement in physical activities, and sexual behaviour.
- Immigrants were not identified by immigrant class, e.g., refugees.
- Some population exclusions:
  - Non tax filers, under the age of 25, institutional residents at cohort inception, those not enumerated by 1991 long form census.
- Ongoing data linkage development at Statistics Canada attempt to address these limitations.
Strengths

- Large size and representative of most population groups (immigrants, Aboriginals).
  - In the current study, has permitted more realistic assessment of mortality differentials by immigrant status.
- Population based.
- Simultaneous analysis of several variables.
  - Multilevel analysis.
- Long latency period required for cancer outcomes.
- Captures residential mobility over a 27 year period.
  - Environmental exposure via the use of postal code representative points.
Conclusions

Question 1: Results point to selection effects:

- Cultural effects - Differences by source countries.
- Canada’s immigration system:
  - ‘Points-based system selects immigrants on characteristics positively associated with health.
  - People selected mostly healthier because of medical screening.
- Unobservable characteristics.

Question 2: Healthy immigrant effect: Immigrants healthier at arrival, but decline occurs over time:

- Early years - difficulties of integration.
- Later years - acculturation.
Conclusions

- Data artefact and ‘Salmon bias’? – Implausible.

- Our knowledge of immigrant health (and other outcomes) will be further deepened from the ongoing data linkage work.
Data access

- Research Data Centres
  - www.statcan.gc.ca/rdc-cdr

- Centre for Data Development and Economic Research
  - For analysis using Longitudinal Worker File
  - www.statcan.gc.ca/cder-cdcre
**My Contact:**

Walter Omariba, PhD  
Health Analysis Division  
Statistics Canada  
Ottawa, ON  
Tel: (613) 853-4067  
Email: alter.omariba@statcan.gc.ca