

## Original quantitative research

# Examining the municipal-level representativeness of the Canadian Longitudinal Study on Aging (CLSA) cohort: an analysis using Calgary participant baseline data

Samantha J. Norberg, MSW (1,2); Ann M. Toohey, PhD (2,3); Sian Jones, MScEcon, MPP (4); Raynell McDonough, MSW (4); David B. Hogan, MD (2,3,5)

This article has been peer reviewed.

 [Tweet this article](#)

### Abstract

**Introduction:** The Canadian Longitudinal Study on Aging (CLSA) is a rich, nationally representative population-based resource that can be used for multiple purposes. Although municipalities may wish to use CLSA data to address local policy needs, how well localized CLSA cohorts reflect municipal populations is unknown. Because Calgary, Alberta, is home to one of 11 CLSA data collection sites, our objective was to explore how well the Calgary CLSA sample represented the general Calgary population on select sociodemographic variables.

**Methods:** Baseline characteristics (i.e. sex, marital status, ethnicity, education, retirement status, income, immigration, internal migration) of CLSA participants who visited the Calgary data collection site between 2011 and 2015 were compared to analogous profiles derived from the 2011 National Household Survey (NHS) and 2016 Census datasets, which spanned the years when data were collected on the CLSA participants.

**Results:** Calgary CLSA participants were representative of the Calgary population for age, sex and Indigenous identity. Discrepancies of over 5% with the NHS and/or 2016 Census were found for marital status, measures of ethnic diversity (i.e. immigrant status, place of birth, non-official language spoken at home), internal migration, income, retirement status and education.

**Conclusion:** Voluntary studies face challenges in recruiting fully representative cohorts. Communities opting to use CLSA data at a municipal level, including the 10 other CLSA data collection sites, should exercise caution when interpreting the results of these analyses, as CLSA participants may not be fully representative of the local population on select characteristics of interest.

**Keywords:** demographics, Calgary, cities, longitudinal studies, census, Canadian Longitudinal Study on Aging, CLSA

### Introduction

Like many other municipalities in Canada and around the world, the City of Calgary, Alberta, implemented an age-friendly strategy in 2015. Based on the World Health

Organization's *Global Age-friendly Cities* guide,<sup>1-3</sup> the strategy's vision states that "Calgary is an age-friendly city where all people have lifelong opportunities to thrive."<sup>3,p.12</sup> To achieve this vision, the City and other stakeholder groups identified

### Highlights

- The Canadian Longitudinal Study on Aging (CLSA) is designed to be nationally and provincially representative for age and sex.
- Municipal representativeness of CLSA data is unknown.
- We compared baseline sociodemographic characteristics of Calgary CLSA participants in the comprehensive cohort with those derived from the 2011 National Household Survey and 2016 Canadian Census.
- Calgary's CLSA sample was representative for age, sex and Indigenous identity but was not fully representative for ethnic diversity, internal migration, education and income when compared to the true population.
- Researchers, planners, policy makers and others using municipal-level CLSA data should consider representativeness of their CLSA sample when interpreting findings.

short- and mid-range actions.<sup>3</sup> The Canadian Longitudinal Study on Aging (CLSA) was recommended as a data resource that could assist in both informing and evaluating the strategy by establishing baseline measures and tracking changes over time.<sup>4</sup>

The CLSA is a national research platform designed to advance our understanding of

### Author references:

1. Faculty of Social Work, University of Calgary, Calgary, Alberta, Canada
2. Brenda Strafford Centre on Aging, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
3. Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
4. Calgary Neighbourhoods, The City of Calgary, Calgary, Alberta, Canada
5. Division of Geriatric Medicine, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

**Correspondence:** Ann M. Toohey, 3280 Hospital Dr NW, Calgary, AB T2N 4Z6; email: amtoohey@ucalgary.ca

the complexities of aging. The overall aims of the CLSA are to examine aging as a dynamic process; investigate the inter-relationship between intrinsic and extrinsic factors, from mid-life to older age; capture the transitions, trajectories and profiles of aging; and provide infrastructure and build capacity for sustained high-quality research on aging.<sup>5</sup> Many of the characteristics being tracked longitudinally were selected to both support population-based research and lead to evidence-informed government policies.<sup>6</sup>

Launched in 2011, the CLSA is a rich data source available to planners and policy makers at all levels of government and academic researchers. While participation is voluntary, the CLSA sampling framework was designed to achieve national and provincial representativeness for age and sex. The extent to which CLSA data are representative for these variables at the municipality level is not known. There are also concerns about its representativeness for other sociodemographic characteristics that may be of interest. These questions merit attention if municipalities use CLSA data to assess population-level needs, monitor sociodemographic changes and understand the impact of public policy implementation.

To understand the extent to which CLSA measures are generalizable to the true population, it was important to assess how closely the CLSA Calgary sample mirrored the sociodemographic profile of middle-aged and older residents of Calgary, Alberta. Other studies have examined the limitations of establishing study cohort representativeness through comparisons with census data.<sup>7-11</sup> Volunteer and selection bias that lead to underrepresentation of minorities and other vulnerable groups may be a possible threat to the generalizability of results derived from volunteer-based cohort studies.<sup>7-11</sup> Examining cohort representativeness can help those using these data to determine when other data sources or analytical approaches should be utilized.

Calgary hosts one of the 11 CLSA data collection sites located across Canada. Assessing the representativeness of Calgary's CLSA sample would inform the City of Calgary administration of the strengths and limitations of CLSA data in evaluating the characteristics and needs of the local older population and the effectiveness of

age-friendly policy implementation. The objective of this current study was to evaluate the extent to which baseline CLSA Calgary municipal data were representative of the corresponding geographical level "true" population as captured by either the 2011 National Household Survey (NHS) or 2016 Census data. This examination would assess the utility of the CLSA as a prospective longitudinal data source for tracking age-friendly policy implementation. Although we focussed on Calgary data to explore representativeness, other cities with a CLSA data collection site could replicate our analysis to evaluate their age-friendly policy implementation.

## Methods

### *Data sources – baseline CLSA data*

The CLSA is a national voluntary study that consists of two cohorts.<sup>5</sup> The tracking cohort is made up of 21 241 randomly selected participants from across Canada who provide alphanumeric data via computer-assisted telephone interviews. The comprehensive cohort consists of 30 097 randomly selected participants living within a 25-km radius of one of 11 data collection sites spread across Canada. These participants provide alphanumeric data, undergo detailed in-person assessments and provide biological samples.

Study baseline data collection for both the tracking and comprehensive cohorts began in 2011 and was completed in 2015. The intent is to follow participants for 20 years or until death, whichever comes first.

Two different sampling designs were used to recruit the two study cohorts. A national sampling frame was used to ensure representativeness for age and sex. In the tracking cohort, people living in postal code areas with average lower education achievement were oversampled to adjust for bias toward recruiting participants with higher socioeconomic status and to ensure sufficient heterogeneity for analyses.<sup>6</sup> Individuals living within a 25-km radius of CLSA data collection sites were intentionally oversampled for inclusion in the comprehensive cohort to receive physical examinations and provide biological samples.<sup>5</sup> For a more detailed description, see Raina et al.<sup>5</sup>

For the tracking cohort, sampling was designed to provide results that would be generalizable at a national level and by

province in relation to the overall age and sex distribution of the population. Three sampling frames were used: the Canadian Community Health Survey on Healthy Aging; provincial health registries (except in the provinces of Alberta and Quebec); and random-digit dialling to landlines (as distinguished from a mobile cellular line).<sup>12</sup> Participants in the comprehensive cohort were recruited using provincial health registries (as before, except for in Alberta and Quebec) and random-digit dialling sampling frames.

The overall response rate was 10% and the participation rate was 45%.<sup>5</sup> Individuals were excluded from participating in the CLSA if they could not speak or write in English or French, were living in an institution, were unable to provide informed consent at the time of enrolment, resided in any of the three Canadian territories, were living on federal First Nations reserves or were full-time members of the Canadian Armed Forces.

We initially examined baseline data of CLSA participants in both the tracking ( $n = 306$ ) and comprehensive ( $n = 2956$ ) cohorts who, at the time of recruitment, lived in Calgary. We examined data for participants aged 45 to 64 years and 65-plus years separately. After assessing the similarities and differences between the cohorts, we elected to utilize only the comprehensive cohort dataset, with its response rate of approximately 11%.<sup>12</sup>

Participants in the tracking cohort were ultimately excluded because some measures of interest to the municipality for evaluating its age-friendly strategy (e.g. life space index) were either not available or were collected in a manner that may have affected the responses (e.g. elder abuse, as tracking cohort participants underwent a telephone interview where other people may have been in hearing distance, whereas comprehensive cohort participants underwent a confidential face-to-face interview).

Finally, we excluded 109 comprehensive cohort participants (3.7% of the sample) whose residential address was outside of the City of Calgary legal jurisdiction.

Of those included in the final sample ( $n = 2847$ ), 1640 were aged 45 to 64 years and 1207 were aged 65-plus years.

## Data sources – 2011 NHS and 2016 Census data

Because baseline CLSA data were collected between 2011 and 2015, we compared CLSA data with both 2011 NHS<sup>13</sup> and 2016 Census<sup>14</sup> data to determine the distribution of characteristics for the “true” population of Calgary.

In 2011, completion of the long-form questionnaire of the Canadian census was not mandatory. The 2011 NHS, collecting similar data to the long-form questionnaire, sampled 3 out of 10 households; participation was voluntary. The NHS included questions from previous iterations of the national census. The present study accessed 2011 NHS data through a public use microdata file released by Statistics Canada, reporting a 25% sample of the collected data. Our approach also reflects the CLSA’s assessment of the national representativeness of the data, which also made comparisons to 2011 NHS data, as the best available representation of the true Canadian population.<sup>5</sup>

The 2016 Census long-form questionnaire sampled 25% of the Canadian population. Mandatory participation had been reinstated in 2015. Due to the timing of our analysis, the 2016 Census public use microdata file for Calgary had not yet been released. Therefore, we accessed 2016 Census data through public aggregated tables reporting on a 25% sample of the collected data. The data tables were accessible using “Beyond 20/20,” a platform that Statistics Canada uses to disseminate aggregate data.

Because the 2011 NHS public use microdata file and the 2016 Census aggregated table data are organized into 5-year age cohorts, we merged available data to create the 45- to 64-year and the 65-plus-year age group categories. From the 2011 NHS, we created a dataset describing the Calgary population aged 45 to 64 years ( $n = 8808$ ) and 65-plus years ( $n = 2849$ ). The datasets created from the 2016 Census described the Calgary population aged 45 to 64 years ( $n = 319\,600$ ) and 65-plus years ( $n = 127\,880$ ).

### Characteristics

We examined sex, marital status, immigration status, place of birth, Indigenous identity, language most often spoken at home, education, working status, total

personal income and internal migration status (defined as a within-Canada household relocation that took place within the past 5 years).

Data for these variables were categorized and presented as percentages. For accurate comparisons between data sources, variable recoding was used to collapse 2011 NHS and 2016 Census response categories as needed. Across all data sources, education, working status and language most spoken at home variables were recoded for comparability. Additional recoding of marital status, country of birth, Indigenous identity and total personal income variables in the 2011 NHS and 2016 Census was done to allow for comparability with CLSA data. Since neither comparison dataset reported a retirement variable, we used the “not in the labour force” measure as a proxy for retirement status. Comparative internal migration variables were also derived from corresponding variables in the CLSA, 2011 NHS and 2016 Census datasets. Finally, we created a category for Canada as the country of birth in the 2011 NHS and 2016 Census datasets using the “non-immigrant” measure.

The CLSA dataset had 5.0% or less missing data (i.e. data coded as “refused,” “required question was not answered,” “at least one required question was not answered” and “don’t know/no answer” responses combined) for the sociodemographic variables in this analysis. The exception was “personal income,” with 7.3% missing data for the 45- to 64-year age group and 12.5% missing data for the 65-plus-year age group. The 2011 NHS variables contained 5% or less missing data for all sociodemographic variables assessed in this analysis. It is important to note that while we excluded observations with missing values from our CLSA sample, the NHS values for missing data were imputed using the nearest-neighbour method described in the *NHS User Guide*.<sup>15</sup> At a national level, response rates for education, income and work items in the NHS questionnaire were lower than for the other characteristics measured, and the values for these items were more likely to be imputed.<sup>15</sup>

### Analysis

Sociodemographic profiles were established for baseline CLSA (collected between 2011 and 2015), 2011 NHS and 2016 Census

cohorts living within Calgary’s jurisdictional boundaries, as defined using Forward Sortation Areas (i.e. the first three digits of participants’ postal codes). We stratified our analysis separately for the 45- to 64-year and 65-plus-year age groups because people become eligible for policy-driven programs like pensions and subsidies and other age-friendly programs and activities upon reaching age 65. For CLSA data, descriptive variables were adjusted for sampling probabilities using the trimmed analytic weights provided by CLSA, which adjust for inclusion probability.<sup>12</sup> The 2016 Census sociodemographic frequencies were calculated using Microsoft Excel, as per the format of the released data. All other analyses were conducted using SPSS version 25.0.<sup>16</sup>

Rather than looking for statistical significance (trivial differences in proportions can be statistically significant in studies with a large number of participants), we were interested in the *practical importance* of any differences seen, that is, whether differences would have real and noticeable effects on the interpretation of the data. We decided a priori on a 10% difference in proportions between the CLSA sample and the true population as a threshold for practical importance to consider when interpreting findings. We also noted percentage differences of 5% to 9% as being of questionable importance when interpreting findings.

## Results

### Baseline CLSA (2011–2015) and 2011 NHS comparisons

Both sex and Indigenous identity distributions in the CLSA sample were representative across both age categories (see Table 1). In contrast, immigrants were underrepresented in the CLSA sample, as indicated by differences in place of birth (and Asia in particular) and language spoken at home, suggesting practical differences compared with 2011 NHS data.

In both age groups, but especially among those aged 65-plus years, lower-income Calgary residents were underrepresented in the CLSA sample. Marital status and educational achievement were representative for Calgary CLSA participants aged 65-plus years. For the 45- to 64-year age group, discrepancies in marital status were questionable, with married individuals overrepresented and divorced individuals

**TABLE 1**  
Comparison of baseline CLSA and 2011 NHS demographic characteristics by age group, Calgary data<sup>a</sup>

Characteristic	45–64 years			≥ 65 years		
	CLSA	2011 NHS	Difference	CLSA	2011 NHS	Difference
Sample size (n) <sup>b</sup>	1640	8808	–	1207	2849	–
<b>Sex (%)</b>						
Male	50.7	50.6	0.1	46.5	44.6	1.9
Female	49.3	49.4	–0.1	53.5	55.4	–1.9
<b>Marital status (%)</b>						
Married/common-law	81.7	73.9	7.8*	68.0	63.1	4.9
Single	7.9	8.8	–0.9	5.0	3.8	1.2
Widowed	2.2	2.4	–0.2	17.0	21.9	–4.9
Divorced	6.0	11.7	–5.7*	8.6	9.6	–1.0
Separated	2.1	3.2	–1.1	1.3	1.6	–0.3
<b>Immigrant status (%)</b>						
Immigrant	16.7	32.2	–15.5**	25.4	39.5	–14.1**
Non-immigrant	83.3	67.8	15.5**	74.6	60.5	14.1**
<b>Place of birth (%)<sup>c</sup></b>						
Canada	83.3	67.8	15.5**	74.6	60.5	14.1**
Other North America	1.5	1.4	0.1	1.7	1.3	0.4
South America, Central America, Caribbean	2.0	2.3	–0.3	1.5	1.6	–0.1
Europe	8.6	9.3	–0.7	19.6	18.9	0.7
Africa	0.6	2.1	–1.5	0.9	1.7	–0.8
Asia	3.2	16.7	–13.5**	1.4	15.4	–14.0**
Oceania and others	0.9	0.4	0.5	0.3	0.6	–0.3
<b>Indigenous identity (%)</b>						
Indigenous	3.7	1.9	1.8	2.3	0.7	1.6
Non-Indigenous	96.3	98.1	–1.8	97.7	99.3	–1.6
<b>Language most spoken at home (%)</b>						
English/French	94.8	86.4	8.4*	97.2	79.8	17.4**
Other	5.2	13.6	–8.4*	2.8	20.2	–17.4**
<b>Postsecondary degree/diploma (%)</b>						
No	27.6	33.7	–6.1*	51.5	49.3	2.2
Yes	72.4	66.3	6.1*	48.5	50.7	–2.2
<b>Working status (%)</b>						
Not retired	86.7	81.8	4.9	26.8	18.4	8.4*
Retired	13.3	16.3	–3.0	73.3	69.7	3.6
Never worked	n/a	2.0	n/a	n/a	11.9	n/a
<b>Total personal income (%)</b>						
< \$20 000	14.3	23.2	–8.9*	16.1	30.6	–14.5**
\$20 000–49 999	21.7	26.8	–5.1*	46.9	44.0	2.9
\$50 000–99 999	35.3	30.3	5.0*	27.4	19.4	8.0*
\$100 000–149 999	15.7	9.4	6.3*	5.5	2.4	3.1
≥ \$150 000	13.0	10.3	2.7	4.1	3.7	0.4
<b>Internal migration status (%)<sup>d</sup></b>						
Non-movers	97.0	92.6	4.4	97.9	95.6	2.3
Moved home within community	1.6	5.4	–3.8	1.6	3.0	–1.4
Moved home and community	1.4	1.3	0.1	0.6	0.9	–0.3

Footnotes on the following page

**TABLE 1 (footnotes)**  
**Comparison of baseline CLSA and 2011 NHS demographic characteristics by age group, Calgary data<sup>a</sup>**

**Abbreviations:** CLSA, Canadian Longitudinal Study on Aging; n/a, not applicable; NHS, National Household Survey.

<sup>a</sup> Calgary data are derived from the CLSA comprehensive sample from the Alberta data collection site, which is located in Calgary. Only participants living within Calgary's jurisdictional region were included in this analysis.

<sup>b</sup> Sample sizes are not weighted. All other proportions were adjusted for sampling probabilities using inflation weights provided by the CLSA, based upon the size of the community-dwelling population living near the data collection site in 2011.

<sup>c</sup> The "Canada" category was manually calculated and added to this variable for comparison purposes.

<sup>d</sup> Defined as a within-Canada household relocation within the past 5 years.

<sup>\*</sup> Questionably important difference (5%–9%).

<sup>\*\*</sup> Practically important difference ( $\geq 10\%$ ).

underrepresented. Differences in education were also questionable. Older CLSA participants ( $\geq 65$  years) were more likely to be not retired.

### **Baseline CLSA (2011–2015) and 2016 Census comparisons**

Baseline CLSA (2011–2015) and 2016 Census data comparisons were similar to the CLSA and 2011 NHS comparisons with just a few exceptions (see Table 2). In the older age category, discrepancies in educational achievements were greater, achieving a practical versus questionable level of importance. For both age groups, differences in the proportions of participants who had moved within Canada during the past 5 years were of practical importance compared with the 2016 Census data.

### **Discussion**

CLSA data can be used to provide valuable insights into the social and physical characteristics of middle-aged and older adults in Canada. In addition to being a key data source for researchers, the CLSA is of potential use for all levels of governments that are trying to understand their aging population and evaluate the impact of age-friendly policies such as "aging-in-place" (i.e. safely remaining in one's community throughout older age).<sup>17,18</sup>

The sampling frame the CLSA employs ensures national and provincial representativeness of the sample by age and sex. At the national level, the CLSA sample, and particularly the comprehensive cohort, is characterized by higher levels of education and household income as well as higher percentages of Canadian-born participants, compared with 2011 Census data.<sup>5</sup> The objective of our study was to explore how well baseline data collected from CLSA participants (between 2011 and 2015) reflected the "true" population at a local municipal level. To do this, we

compared these data with both the 2011 NHS and 2016 Census data. Establishing a baseline description of the population would also be helpful when using CLSA and other data sources to evaluate policy implementation over time, given the longitudinal design of the CLSA.

In general, we found the largest discrepancies between CLSA and 2016 Census estimates. While Calgary CLSA participants were representative for sex and Indigenous identity, marital status, ethnic diversity (i.e. immigrant status, place of birth and language spoken at home), education, working status, personal income and internal migration diverged from the true population. For some of these measures, the differences between CLSA and NHS/Census estimates were greater than 10%, which we viewed as of practical importance when using CLSA data to assess the local population. These findings suggest that CLSA data for municipally defined populations may underrepresent certain marginalized populations, which could affect interpretation. The importance of accounting for these differences will depend on the research questions being asked and how the results will be used.

We used Calgary as a case study for evaluating the representativeness of CLSA data at a municipal level. While we cannot comment on the representativeness of CLSA data for other municipalities, our findings may inform others who are considering using CLSA data to describe the health and well-being of their aging population to assess the impact of public policy.<sup>19,20</sup> Calgary differs from other Canadian cities in terms of its economic and socio-demographic profile.<sup>21</sup> Others are advised to conduct similar comparisons for their own local setting. This would be particularly applicable for other cities with CLSA data collection sites (i.e. Surrey, BC; Vancouver, BC; Victoria, BC; Winnipeg, MB; Hamilton, ON; Ottawa, ON; Montréal,

QC; Sherbrooke, QC; Halifax, NS; and St. John's, NL). Although the CLSA was not designed to be representative at a municipal level, there is growing interest in using the data to address local questions.<sup>21</sup>

The CLSA's inclusion and exclusion criteria may partially account for some of the differences observed. Because CLSA participants were required to be fluent in English or French, this may have systematically eliminated some older people who had been born outside of Canada. It is unclear, however, why people who reported being born in Asia were particularly underrepresented in our setting. And while efforts were made to over-sample CLSA participants with lower levels of education,<sup>5,6</sup> we found that participants with lower-level education attainment were nonetheless underrepresented.

Because of provincial privacy legislation, sampling from provincial health registries was not permitted in Alberta or Quebec.<sup>12</sup> Calgary participants were exclusively recruited using random-digit dialling of landlines.<sup>12</sup> This use of landlines may help explain the overrepresentation of those who had not moved in the past 5 years. The proportion of Canadian households with a landline has diminished over the last few years.<sup>22</sup> In addition, it is likely that someone who has recently moved would be more inclined to only use mobile telephone services in their new community,<sup>22</sup> and at the time of baseline data collection, Calgary was experiencing a high level of net internal migration.<sup>23</sup>

The random-digit dialling approach may also help account for the underrepresentation of lower-income households; these households are more likely to report using mobile telephone services only and no landline.<sup>22</sup> In Australia, where telecommunications trends are similar to those in Canada, Barr et al. found that relying on landlines only when sampling for a cohort

**TABLE 2**  
**Comparison of baseline CLSA and 2016 Census demographic characteristics by age group for Calgary data<sup>a</sup>**

Characteristic	45–64 years			≥ 65 years		
	CLSA	2016 Census	Difference	CLSA	2016 Census	Difference
Sample size (n) <sup>b</sup>	1640	319 600	–	1207	127 880	–
<b>Sex (%)</b>						
Male	50.7	50.0	0.7	46.5	46.8	–0.3
Female	49.3	50.0	–0.7	53.5	53.2	0.3
<b>Marital status (%)</b>						
Married/common-law	81.7	73.1	8.6*	68.0	64.1	3.9
Single	7.9	10.6	–2.7	5.0	4.1	0.9
Widowed	2.2	2.2	0.0	17.0	18.8	–1.8
Divorced	6.0	10.8	–4.8	8.6	11.2	–2.6
Separated	2.1	3.3	–1.2	1.3	1.9	–0.6
<b>Immigrant status (%)<sup>c</sup></b>						
Immigrant	16.7	38.1	–21.4**	25.4	43.6	–18.2**
Non-immigrant	83.3	61.0	22.3**	74.6	56.0	18.6**
<b>Place of birth (%)<sup>d</sup></b>						
Canada	83.3	61.6	21.7**	74.6	56.2	18.4**
Other North America	1.5	1.2	0.3	1.7	1.4	0.3
South America, Central America, Caribbean	2.0	3.2	–1.2	1.5	2.5	–1.0
Europe	8.6	7.9	0.7	19.6	17.3	2.3
Africa	0.6	3.3	–2.7	0.9	2.4	–1.5
Asia	3.2	22.4	–19.2**	1.4	19.8	–18.4**
Oceania and others	0.9	0.4	0.5	0.3	0.5	–0.2
<b>Indigenous identity (%)</b>						
Indigenous	3.7	2.3	1.4	2.3	1.3	1.0
Non-Indigenous	96.3	97.7	–1.4	97.7	98.7	–1.0
<b>Language most spoken at home (%)</b>						
English/French	94.8	82.5	12.3**	97.2	79.5	17.7**
Other language	5.2	17.5	–12.3**	2.8	20.5	–17.7**
<b>Postsecondary degree/diploma (%)</b>						
No	27.6	33.6	–6.0*	23.0	47.8	–24.8**
Yes	72.4	66.4	6.0*	77.0	52.2	24.8**
<b>Working status (%)</b>						
Not retired	86.7	80.6	6.1*	26.8	20.2	6.6*
Retired	13.3	19.4	–6.1*	73.3	79.8	–6.5*
<b>Total personal income (%)</b>						
< \$20 000	14.3	18.8	–4.5	16.1	26.3	–10.2**
\$20 000–49 999	21.7	26.2	–4.5	46.9	43.8	3.1
\$50 000–99 999	35.3	31.0	4.3	27.4	21.6	5.8*
\$100 000–149 999	15.7	11.9	3.8	5.5	4.1	1.4
≥ \$150 000	13.0	12.1	0.9	4.1	4.2	–0.1
<b>Internal migration status (%)<sup>e</sup></b>						
Non-movers	97.0	70.0	27.0**	97.9	80.3	17.6**
Moved home within community	1.6	20.9	–19.3**	1.6	13.1	–11.5**
Moved home and community	1.4	9.0	–7.6*	0.6	6.6	–6.0*

Footnotes on the following page

**TABLE 2 (footnotes)**  
**Comparison of baseline CLSA and 2016 Census demographic characteristics by age group for Calgary data<sup>a</sup>**

**Abbreviations:** CLSA, Canadian Longitudinal Study on Aging; NHS, National Household Survey.

<sup>a</sup> Calgary data are derived from the CLSA Comprehensive sample from the Alberta data collection site, which is located in Calgary. Only participants living within Calgary's jurisdictional region were included in this analysis.

<sup>b</sup> Sample sizes are not weighted. All other proportions were adjusted for sampling probabilities using inflation weights provided by the CLSA, based upon the size of the community-dwelling population living near the data collection site in 2011.

<sup>c</sup> Percentages for 2016 data do not add up to 100% as we do not include a "non-permanent resident" category (not included in the CLSA).

<sup>d</sup> The "Canada" category was manually calculated and added to this variable for comparison purposes.

<sup>e</sup> Defined as a within-Canada household relocation within the past 5 years.

<sup>\*</sup> Questionably important difference (5%–9%).

<sup>\*\*</sup> Practically important difference (≥ 10%).

study reduced the accuracy of estimates for certain health indicators compared with when using a combination of landline and mobile phones.<sup>24</sup>

Our decision to remove CLSA participants who lived outside of the City of Calgary jurisdiction from our analysis may have also influenced the extent to which the representativeness of the sampling frame was maintained.

### **Strengths**

Strengths of this study include the relative novelty of assessing the representativeness of a municipally defined subsample of CLSA participants on characteristics other than age and sex. A similar methodological approach has been applied to comparable data sources in France,<sup>7</sup> Australia<sup>9</sup> and the USA,<sup>10</sup> and more recently in Canada.<sup>8,11</sup>

The CLSA research team has explored and reported on the extent to which the CLSA cohort is representative, at a national level, on select sociodemographic characteristics.<sup>5</sup> This current research investigates this question at the municipal level, addressing concerns that representativeness at the municipal level of analysis might differ from that found at the national level. Our endeavour is relevant, as interest in using CLSA data for municipally driven age-friendly initiatives is growing, and those using CLSA data to evaluate, for example, the impact of age-friendly policy, should be aware that some local subgroups may not be well represented. Identifying discrepancies between the characteristics of CLSA participants and the "true" population, coupled with thoughtful judgements about the importance of these discrepancies in understanding the state of the older local population, will help guide policy makers in applying their findings.

Our current study also contributes to the growing interest in population health intervention research methodologies focusing on activities that lie outside of traditional public health or health care jurisdictions that influence the health and well-being of the population.<sup>25</sup> As the CLSA begins to release longitudinal data, understanding the generalizability of the data will equip policy makers and academic researchers to make more informed interpretations of the potential connection between policy implementation and population health.

### **Limitations**

Our study offers a model that others can adapt to their own municipal contexts. Nevertheless, there are also a number of limitations to consider. For instance, we made judgements based on what we considered to be practically or questionably important differences between CLSA and 2011 NHS or 2016 Census data. There is no consistent precedent for these judgements in the literature.

Another limitation is the exclusion of missing CLSA data from this analysis. For most variables, the proportion was nominal, but for one (i.e. "personal income") the proportion was above the desired threshold of 5% that we had established. Although high rates of missing income data are not unusual in population-based studies, this leads to a less precise categorization of the variable and possibly a degree of bias if the missing values are clustered within particular social subgroups.<sup>26</sup> However, the consistency of the trends in several variables (i.e. retirement and education in relation to income; immigration status and place of birth in relation to ethnic diversity) affirms our cautious conclusions regarding the representativeness of the CLSA sample. As already noted, the CLSA was not designed

to provide representative data at a municipal level.

It is also important to note limitations with the census data sources that we used as proxies for the "true" population. We used Canadian population data from both 2011 and 2016 as comparison data because these straddled the period when CLSA baseline data were collected. Changes in federal government policies to do with privacy led to permitting voluntary versus mandatory completion of portions of Canada's 2011 Census. This decision was reversed in time for the 2016 Census.<sup>27</sup> Although the national weighted response rate was 77.2% for the NHS,<sup>15</sup> the voluntary nature of participation may have introduced systematic biases into this dataset. Future waves of CLSA data should be compared with 2016 and subsequent census datasets, where participation is mandatory and the true population is better reflected.

The public use microdata file used to access the 2011 NHS data and the "Beyond 20/20" aggregated tables used to access the 2016 Census data report on just 25% of the data collected. Most municipalities will have access to these publicly available forms of census data in conducting their evaluation activities.

Finally, recoding variables contained within these data sources in order to make meaningful comparisons with CLSA data posed methodological challenges. These included merging age groups, as was necessary for both the 2011 and 2016 Census data. We also note that our decision to use "Not in the labour force" to indicate retirement may have led to an inflated number of people considered retired in both datasets. However, these data are counterbalanced by comparisons of those reporting being "not retired" in the CLSA and those

reporting being in the labour force in both the 2011 NHS and 2016 Census questionnaires.

Our analysis provides a case study that focuses on a single municipality. Exploring the representativeness in other municipalities with a CLSA data collection site was beyond the scope of our research objectives, data application agreement for the use of CLSA data and institutional ethics approval. We intended that our work support the City of Calgary in their exploration of data sources that could be used in the evaluation of their age-friendly strategy.<sup>3</sup>

## Conclusion

The CLSA is a valuable national resource for extending our understanding of determinants of health and well-being later in life, yet voluntary studies face challenges in recruiting representative cohorts. Our study examined the extent to which the baseline CLSA sample is representative of the actual Canadian population at the municipal level, specifically for Calgary, Alberta. Notable differences in sociodemographic characteristics were observed between the CLSA subsample of Calgary participants surveyed between 2011 and 2015 and true comparable populations as described by 2011 NHS and 2016 Census data. Ethnic diversity was underrepresented within Calgary's CLSA subsample, as were older participants reporting lower education and personal income.

Researchers and policy makers who use this important dataset to explore locally defined populations should also be aware of the potential sampling limitations. We recommend that municipalities that utilize CLSA data at a local level conduct a similar analysis to the one we performed, comparing geographically defined CLSA data with local or national census data. The use of multiple data sources is also recommended in order to triangulate or enrich interpretation of the findings obtained, especially those with implications for traditionally underserved populations such as lower socioeconomic status and ethnically diverse older adults.

## Ethics approval

REB approval obtained (University of Calgary REB17-0692).

## Acknowledgements

This analysis was made possible using the data collected by the Canadian

Longitudinal Study on Aging (CLSA). The Government of Canada funded the CLSA through the Canadian Institutes of Health Research (grant reference LSA 9447) and the Canada Foundation for Innovation. Samantha Norberg was funded through an Alberta Health Services "Seniors Health Strategic Clinical Network Summer Studentship" award and by the Brenda Strafford Centre on Aging at the University of Calgary.

The authors also wish to thank Peter Peller, Tyler Williamson, Sohel Nazmul and Lauren Griffith for their valuable assistance. The authors also wish to thank two anonymous reviewers and the associate editor of this journal for their contributions.

## Conflicts of interest

Dr. David B. Hogan is the Local Responsible Investigator for the Calgary data collection site of the Canadian Longitudinal Study on Aging.

## Authors' contributions and statement

AMT and DBH developed the research objectives and research design. SJN and SJ led the analysis of the data. SJN and AMT led the drafting and revising of the manuscript. All authors contributed to interpreting the data and revising the manuscript. All authors approved the manuscript for publication.

The content and views expressed in this article are those of the authors and do not necessarily reflect those of the Government of Canada.

## References

1. World Health Organization. Global age-friendly cities: a guide. Geneva (CH): WHO; 2007. [https://www.who.int/ageing/publications/Global\\_age\\_friendly\\_cities\\_Guide\\_English.pdf](https://www.who.int/ageing/publications/Global_age_friendly_cities_Guide_English.pdf)
2. Public Health Agency of Canada. Age-friendly communities [Internet]. Ottawa (ON): Government of Canada; 2009 [cited 2018 Sep 24]. Available from: <https://www.canada.ca/en/public-health/services/health-promotion/aging-seniors/friendly-communities.html#sec3>

3. City of Calgary. Seniors age-friendly strategy and implementation plan 2015-2018 [Internet]. Calgary (AB): City of Calgary, Community & Neighbourhood Services; 2015 [cited 2016 Mar 31]. Available from: [http://www.calgary.ca/CSPS/CNS/Documents/seniors/Seniors\\_Age\\_Friendly\\_Strategy.pdf](http://www.calgary.ca/CSPS/CNS/Documents/seniors/Seniors_Age_Friendly_Strategy.pdf)
4. Tunstall LA, Eaton SE. City of Calgary seniors age-friendly strategy (SAFS) evaluation report phase 2: determination and selection of population-level indicators. Calgary (AB): City of Calgary; 2016.
5. Raina P, Wolfson C, Kirkland S, et al. Cohort profile: the Canadian Longitudinal Study on Aging (CLSA). *Int J Epidemiol*. 2019;48(6):1752-53j. <https://doi.org/10.1093/ije/dyz173>
6. Kirkland SA, Griffith LE, Menec V, et al. Mining a unique Canadian resource: the Canadian Longitudinal Study on Aging. *Can J Aging*. 2015; 34(3):366-77. <https://doi.org/10.1017/S071498081500029X>
7. Andreeva VA, Salanave B, Castetbon K, et al. Comparison of the sociodemographic characteristics of the large NutriNet-Santé e-cohort with French Census data: the issue of volunteer bias revisited. *J Epidemiol Community Health*. 2015;69(9):893-8. <https://doi.org/10.1136/jech-2014-205263>
8. Dummer TJ, Awadalla P, Boileau C, et al. The Canadian Partnership for Tomorrow Project: a pan-Canadian platform for research on chronic disease prevention. *CMAJ*. 2018; 190(23):E710-7. <https://doi.org/10.1503/cmaj.170292>
9. Hassett A, George K, Harrigan S. Admissions of elderly patients from English-speaking and non-English-speaking backgrounds to an inpatient psychogeriatric unit. *Aust N Z J Psychiatry*. 1999;33(4):576-82. <https://doi.org/10.1080/j.1440-1614.1999.00537.x>
10. Koenig C, Langer-Gould AM, Gould MK, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. *Perm J*. 2012;16(3):37-41. <https://doi.org/10.7812/TPP/12-031>



11. Queenan JA, Williamson T, Khan S, et al. Representativeness of patients and providers in the Canadian Primary Care Sentinel Surveillance Network: a cross-sectional study. *CMAJ Open*. 2016;4(1):E28-32. <https://doi.org/10.9778/cmajo.20140128>
12. Canadian Longitudinal Study on Aging. CLSA technical document: sampling and computation of response rates and sample weights for the tracking (telephone interview) participants and comprehensive participants [Internet]. Hamilton (ON): CLSA; 2011 [cited 2018 Sep 28]. Available from: <https://www.clsa-elcv.ca/doc/1041>
13. Statistics Canada. Individuals File, 2011 National Household Survey (public use microdata files) [Internet]. Ottawa (ON): Statistics Canada; 2011 [cited 2018 Jun 22]. Available from: <https://www150.statcan.gc.ca/n1/en/catalogue/99M0001X>
14. Statistics Canada. Age and sex, and type of dwelling data: key results from the 2016 Census [Internet]. Ottawa (ON): Statistics Canada; 2017 [cited 2017 Jul 2]. Available from: <http://www.statcan.gc.ca/daily-quotidien/170503/dq170503a-eng.htm?HPA=1>
15. Statistics Canada. NHS user guide: National Household Survey, 2011 (Catalogue no. 99-001-X2011001). Ottawa (ON): Government of Canada; 2011.
16. IBM Corporation. IBM SPSS Statistics for Windows, version 25.0. Chicago (IL): IBM Corporation; 2017.
17. Menec VH, Means R, Keating N, Parkhurst G, Eales J. Conceptualizing age-friendly communities. *Can J Aging*. 2011;30(3):479-93. <https://doi.org/10.1017/S0714980811000237>
18. Plouffe LA, Kalache A. Making communities age friendly: state and municipal initiatives in Canada and other countries. *Gac Sanit*. 2011;25:131-7. <https://doi.org/10.1016/j.gaceta.2011.11.001>
19. World Health Organization. Measuring the age-friendliness of cities: a guide to using core indicators. Kobe (JP): World Health Organization; 2015.
20. Plouffe LA, Garon S, Brownoff J, et al. Advancing age-friendly communities in Canada. *Can Rev Soc Policy*. 2012/2013;(68/69):24-38.
21. Baxter D, Enns A, Kristjansson E; Age-Friendly Cities Team. A tale of eight cities: general report on the age-friendliness of eight major Canadian cities. Ottawa (ON): University of Ottawa; 2018.
22. Canadian Radio-television and Communications Commission. Canada's communication system: an overview for Canadians. Section 2.0 [Internet]. Ottawa (ON): CRTC; 2016 [cited 2020 Apr 26]. Available from: <https://crtc.gc.ca/eng/publications/reports/policymonitoring/2016/cmr2.pdf>
23. Saunders D. Report on the demographic situation in Canada: internal migration: overview, 2015/2016 [Internet]. Ottawa (ON): Statistics Canada; 2018 [cited 2020 Aug 28]. [Statistics Canada, Catalogue no. 91-209-X]. Available from: <https://www150.statcan.gc.ca/n1/pub/91-209-x/2018001/article/54958-eng.htm>
24. Barr ML, Ferguson RA, Steel DG. Inclusion of mobile telephone numbers into an ongoing population health survey in New South Wales, Australia, using an overlapping dual-frame design: Impact on the time series. *BMC Res Notes*. 2014;7(1):517. <https://doi.org/10.1186/1756-0500-7-517>
25. Hawe P, Potvin L. What is population health intervention research? *Can J Public Health*. 2009;100(1):Suppl I8-14. <https://doi.org/10.1007/BF03405503>
26. Kim S, Egerter S, Cubbin C, Takahashi ER, Braveman P. Potential implications of missing income data in population-based surveys: an example from a postpartum survey in California. *Public Health Rep*. 2007;122(6):753-63. <https://doi.org/10.1177/003335490712200607>
27. Kilpatrick S. New law ends threat of jailing census resisters [Internet]. Toronto (ON): The Globe and Mail; 2018 Jan 7 [cited 2018 Oct 1]. Available from: <https://www.theglobeandmail.com/news/politics/new-law-ends-threat-of-jailing-census-resisters/article37521191/>