Impact of a change in the calibration strategy for the group of persons aged 65 and over for the Canadian Community health Survey (CCHS)

STATISTICS CANADA - WINTER 2019

Sanderson, Emily – SSMD/DMSS

MCMASTER MATH & STATS – STATS SUBPLAN CO-OP III | SUPERVISOR – RONALD JEAN PAUL

Executive Summary

The Canadian Community Health Survey is used to understand the health needs of Canadians. Estimations are provided for the following age groups: 12-17, 18-34, 35-49, 50-64 and 65+. The 65+ age group has been growing over the past few years, as the large Baby Boomer generation is entering retirement and as life expectancy continues to rise. The need for more detailed information about the 65+ age group is necessary to accurately estimate the needs of the senior population.

The purpose of this study is to find the best way to subdivide the 65+ age group so the users of the Canadian Community Health Survey can make better decisions regarding a population that increasingly depends on health care as they age. This required finding a set of age groups that would have sufficient respondents to be able to run all the programs that are run for the old age groups. Sufficiency is based on a few factors, but in general any combination of Health Region, age group and sex must have at least 20 respondents to be considered reliable and for calibration to run without errors.

In order to test new age groups, they must be run through the normal programs that are being used in production. The program that introduces the age groups was the calibration program that created person level weights using population totals of each age group by Health Region. The first issue was that there lacked population totals for the new age groups by Health Region so these population estimates had to be calculated using available data. Once the population estimates were created for each new age group, the program had to be adjusted to calibrate weights on the new age groups.

The creation of the age groups 65-74 and 75+ was found to be the best division of the age group 65+ as it allowed for enough respondents to properly run calibration while being specific enough to do analysis on this population. The weights produced for these age groups were compared to those calculated in production. This comparison is normally done from year to year in production, using various Health Indicators to calculate significant differences in weights. There were few significant differences between the new and original weights for the new age groups, for every combination of Health Indicator, Health Region and sex. The coefficients of variation of the new estimates were also just as good as those produced from production estimates. This indicated that this new method had the same level of accuracy as previously accepted.

The other age groups that were investigated had insufficient respondents to confidently allow these scenarios to be produced. Grouping the age groups into five year groups was too narrow and the age group 85+ had less than ten respondents for almost every combination of health region, age group and sex so those scenarios had to be rejected.

Table of Contents

Introduction

Canada's senior population is increasing year after year as more Baby Boomers enter retirement so comes a greater need for information to better understand this large, changing population. Currently, Statistics Canada's Canadian Community Health Survey (CCHS) is being used to collect information related to health status, health care utilization and health determinants to help decision makers best allocate funding for health related needs. Currently, all those 65 and older are in one age group which causes issues as this age group contains a large amount of people with highly varied health needs. In order to make better decisions about this diverse population, subdividing the 65 and over age group into smaller age groups will allow us to make better health related decisions about a population that's needs greatly change as they age.

We are interested in researching this topic by Health Region (HR) and Local Health Integration Networks in Ontario (LHIN) which are defined by a region's similar characteristics, such as dividing rural from urban regions.

The main objective of this project is to measure the feasibility and the impact of a change in the calibration strategy for the age group 65 and over for the Canadian Community Health Survey.

We will be interested in answering, if possible, the following question:

Is the sample size for the new age groups that are obtained by splitting the 65 and over category large enough to have reliable estimates of some key variables (e.g. proportion of people with Diabetes)? This will help the users of the CCHS to identify regions that need more or less resources pertaining to specific health determinants.

This report will go through several components of the analysis, starting with the creation of the new age groups and their accompanying datasets, the processes that were adapted and run with the new age groups and the comparison of the results of those processes and the results created in the original production. Finally, a conclusion stating the results of these comparisons and any further research that could be done to further this investigation into creating new age groups.

Creating the Datasets

The first step is to decide on how to divide the 65 plus age group. With the consultation of supervisors and clients, the following five scenarios for new age groups were suggested:

Age Group Scenarios

1 – Two new groupings: 65-69, 70+
2 – Three new groupings: 65-69, 70-74, 75+
3 – Ten year age group: 65-74, 75+
4 – Two ten year age groups: 65-74, 75-84, 85+
5 – Twenty year age group: 65-85, 85+

Using the respondents of the Canadian Community Health Survey we can decide whether all these new age groups have enough respondents to make reliable conclusions.

The next step involves creating the input files needed for the calibration program. These files have the population counts for every age group by HR and LHIN. Unfortunately, at the moment, there does not exist files that indicate the population counts for each of the new age group by HR or LHIN. In order to make up for this, methods to estimate the population counts for the new age groups are needed to continue. In the future, if this project is proven successful, the true value of the population counts would be provided for the new age groups. These files are provided for every month of the year and averages are calculated to produce quarterly and annual population totals.

In order to calculate the population sizes for each new age group, 3 methods were derived using the available information. Population counts by each age and province/territory were used to make the ratios and summations needed in the 3 methods.

Notation

$s : sex = Male\ or\ Female$

$i : \ Lower\ bound\ of\ an\ age\ group$

$j : \ Upper\ bound\ of\ an\ age\ group\ (max\ is\ equivalent\ \ to\ oldest\ age\ of\ certain\ region)$

$k : HR\ or\ LHIN$

$l : Province\ or\ Territory$

$m : year\ (used\ in\ method\ 3)$

$a_{s,i,j,m}$: $Population\ total\ for\ those\ in\ age\ group\ [i,j]\ calculated\ at\ the\ provincial\ or$ $territorial\ level\ for\ a\ given\ month\ in\ a\ given\ year. m\ is\ used\ in\ method\ 3\ only.$

$b_{s,i,j,m}$: $Population\ total\ for\ those\ in\ age\ group\ [i,j]\ calculated\ at\ the\ HR\ or\ LHIN\ level$ $for\ a\ given\ month\ in\ a\ given\ year. m\ is\ used\ in\ method\ 3\ only.$

$x_{s,i,j,l}$: $Ratio\ of\ those\ in\ [i,j]\ over\ all\ those\ 65\ and\ over\ by\ l\ and\ by\ s$

$y_{s,k,l}$: $Ratio\ of\ those\ 65\ and\ older\ of\ a\ specific\ health\ region\ k\ over\ the\ total$ $population\ of\ the\ province\ l\ by\ sex$

$\Delta_{s,m}$: $The\ change\ in\ population\ sizes\ from\ year\ m\ to\ present\ (used\ in\ method\ 3)\ by\ sex$

Method 1

– Calculated the ratio of the population counts of each new age group in each province/territory over the provincial/territorial counts of those 65 and older by sex. For example, the number of females 65-74 over all females those 65 and over. Multiply this ratio by the HR or LHIN population counts of the original age group of 65+ females to get an estimate of each new age group by HR or LHIN. Each variable is rounded to the closest one and below are the formulas used to create the estimates for scenario 2:

$$x_{s,i,j,l} = {}^{a_{s,i,j}}/_{a_{s,65,max}}$$

To obtain the HR population totals for the age groups 65-69, 70-74 and 75+ by sex:

$$b_{s,65,69} = b_{s,65,max} \cdot x_{s,65,69,l}$$

$$b_{s,70,74} = b_{s,65,max} \cdot x_{s,70,74,l}$$

$$b_{s,75,max} = b_{s,65,max} \cdot x_{s,75,max,l}$$

The assumption used in this method is that each HR/LHIN has the same ratio $x_{s,i,j}$ to the province/territory in which it is located. This assumption can only be known to be true in the case of Prince Edward Island, Nunavut, North-West Territories and the Yukon as there exists only one HR in each of these territories and province.

Method 2 – Calculated the proportion of people 65+ in each HR over the sum of all people in that province/territory by sex. In the case of Prince Edward Island, Nunavut, North-West Territories and the Yukon, this value is 1 or 100% due to the fact there is only 1 HR in each of these territories and province. Multiply this proportion by each new age group calculated using the same provincial totals that were used to calculate the estimates of Method 1.

$$y_{s,k,l} = {b_{s,65,max,k}}\Big/{b_{s,min,max,l}}$$

$$\sum_{k=1}^{k=total} b_{s,min,max,k} = b_{s,min,max,l} = a_{s,min,max}$$

To obtain the HR population totals for the age groups 65-69, 70-74 and 75+ by sex:

$$b_{s,65,69} = y_{s,k,l} \cdot a_{s,65,69}$$

$$b_{s,70,74} = y_{s,k,l} \cdot a_{s,70,74}$$

$$b_{s,75,max} = y_{s,k,l} \cdot a_{s,75,max}$$

The assumption used in this method is that each HR (*k*) has proportionally the same number of people in each new age group to its provincial/territorial total population by sex. Method 1 and 2 should be similar for Prince Edward Island and the Territories.

Method 3 - To find the HR population totals for scenario 1, we look at the same cohort five years ago (ex 60-64 age group in 2012 to 2017's 65-69 age group) and calculate the cohort population change using provincial data in the five years. The formula uses the HR datasets from both years and the newly calculated change in population size.

The following formula is calculated by sex and by month with testing year 2017:

$$\Delta_{s,2012} = (a_{s,45,64,2012} - a_{s,50,69,2017}) / (a_{s,45,64,2012})$$

$$b_{s,65,69,2017} = b_{s,45,64,2012} - b_{s,50,64,2017} - \Delta_{s,2012} \cdot b_{s,45,64,2012}$$

$$b_{s,70,max,2017} = b_{s,65,max,2017} - b_{s,65,69,2017}$$

For the age group 70-74, using the 65+ age group from 2007 minus the change in cohort size based on provincial estimates equals the age group 75+ in 2017.
Using 2007 data:

$$\Delta_{s,2007} = (a_{s,65,max,2007} - a_{s,75,max,2017})/(a_{s,65,max,2007})$$

$$b_{s,75,max,2017} = b_{s,65,max,2007} - \Delta_{s,2007} \cdot b_{s,65,max,2007}$$

$$b_{s,70,75,2017} = b_{s,70,max,2017} - b_{s,75,max,2017}$$

This results in all the age groups needed for scenario 1, 2 and 3. Scenario 4 and 5 require going back 20 years and that data is not available for these specific calculations as the CCHS started in 2001. The assumptions used in method 3 is that the percentage change in population size for each HR is the same as the change in population size of the province in which each HR lies. Also that HR with the same names have the same boundaries and characteristics over the 10 years. Proof of the formulas used in method 3 can be found in Appendix A.

It was decided to remove method 2 from further testing as it produced very similar results to method 1 and method 1 was deemed as the stronger option of the two. The ratio used in method 2 was calculated using provincial counts of men or women of all ages. If we decided to improve this method, we would only use provincial counts of those 65 and older to ensure the ratio is the best representation of the target population. In making this improvement, the formula for method 2 became nearly identical to method 1 calculated in a different order, but still producing the same results, with slight differences due to rounding and population estimates.

Method 3 was removed from the calculations as it was unable to produce results for every Health Region. The reason for this is because the HRs occasionally change name and geography over the 5 and ten years so that leaves less data to work with, which was not the goal of this project. There was also the issue of extreme population values, which resulted in negative values. This would have to be resolved with imputing but that degenerates the data even further.

After these conclusions, it was decided to use method 1 datasets in the calibration and bootstrap programs for the scenarios.

Calibration

The CCHS collects data from over 55,000 respondents that are selected to be a representative sample of the entire Canadian population. Calibration methods consist of reweighting units so that survey estimates of totals (or counts or percentages) coincide with true, known population totals or counts or percentages (also called benchmarks) from external sources. Calibration uses auxiliary information as a set of constraints to improve survey estimates by creating "calibrated weights, which are as close as possible, according to a given distance measure, to the original sampling design weights π_k^{-1} while also respecting a set of constraints, the calibration equations" (Deville & Sarndal, 1992).

Before running calibration, person level weights are adjusted to take into account things like nonresponse, out-of-scope, etc. Calibration adjusts these weights even more based on population totals of HR, age group and sex so the sum of all weights of a specific region equals the population total that was previously calculated. This is done using Statistic Canada's Generalized Estimation System, or G-Est. The calibration is performed on HR, age group, sex, province and collection period, and additionally on LHIN for Ontario. There are 4 collection periods per year in the years that were analyzed. One thing to note is a tolerance of 0.5% is allowed at the LHIN level due to the fact the program cannot calibrate exactly on HR, Age Group, Sex and LHIN in Ontario so a little room for error between the sum of the weights and the total population is created due to the fact that the boundaries for LHIN do not match up perfectly to HR boundaries. See Appendix B for a drawing of Ontario HR and LHIN.

When working on scenario 2 which has two additional age groups, the calibration program takes more time to adjust the respondent weights to the population totals. For Ontario, the program also calibrates on LHIN, which causes the calibration to not converge and an error to be produced. In past years an error of 0.5% was tolerated, but in order for the program to work for scenario 2, the tolerance had to be raised to 15.2%, creating a greater difference between the weights and the population totals. Since this only occurs in Ontario it would be up to the CCHS managers and clients to decide what to take from this information and to do further investigation. This also occurred for the population of 2016 with scenario 1 and 2 so we cannot run normal production for those age groups but perhaps if the data users are primarily interested in Health Regions than we can calibrate solely on HR and period in Ontario and eliminate the extra steps. Since the goal of this project is to find an age group that will be feasible in production, these scenarios are rejected.

Another required step for producing estimates is to produce the bootstrap weights. Hall (2003) describes a bootstrap as: "If one defines (as I believe one should) a bootstrap estimator to be the result of replacing an unknown distribution function in the definition of a parameter by its empirical counterpart, then the sample mean is the bootstrap estimator of the population mean."

The first step is to "draw with replacement of B samples (to be called bootstrap replicates) from the parent sample. […] [To] obtain B realizations of *n* independent random variables identically distributed according to [the distribution] $\hat{F}$ amounts to drawing with replacement B samples from the *n* values $\{x_1, x_2, \dots, x_n\}$ of the master sample underlying $\hat{F}$. These B samples are commonly called bootstrap replicates" (Girard, 2009). The next step is to form the bootstrap estimates, for example the mean of each replicate, and then to compute the variance of these estimates. The bootstrap procedure is applied in survey methodology to estimate the variance of calibrated weights. These are called bootstrap weights.

The simplest algebraic expression to create the bootstrap weights is as follows from *A companion to variance estimation written in asymptotically layman terms* (Claude Girard, 2009):

$$w_b(k) = w(k)\frac{n_h}{n_h - 1} mult_b(k)$$

$$mult_b(k) \text{ is the multiplicity of unit } k \text{ in replicate } b$$

$$w(k) \text{ is the original weight}$$

$$w_b(k) \text{is the } b^{th} \text{ bootstrap weight assigned to unit } k$$

$$n_h \text{ is the number of elements in the strata } h$$

If a unit is not selected to be in b, $mult_b(k) = 0$ and then $w_b(k)=0$. In this project, each strata *h* is a combination of Health Region x Age Group (Gage) x Sex and additionally by LHIN in Ontario.

It is necessary that each strata have at least 2 respondents so that if a certain strata is selected to be in a replicate, $n_h - 1$ will be greater than zero. $n_h - 1$ cannot be equal to 0 as it is in the denominator of the bootstrap weight formula. If n_h is equal to 0, the bootstrap weight will equal 0, which will also lead to inconsistencies in the bootstrap replicate totals. If a strata *h* has less than 20 respondents, there runs the risk of a replicate not having at least one respondent for every strata. In this project, bootstrapping works by randomly sampling with replacement 1000 times from the over 55 thousand person level weights. Then the variance is calculated from each of these 1000 replicates to get the variance of the entire dataset.

The clients suggested creating estimates for the age groups found in scenario 4 and 5 but while investigating whether the new age groups had sufficient respondents, it was noted that there were less than 20 respondents for almost every combination of HR, age group, and sex for those 85 and older as shown in Table 1. It was decided to remove scenario 4 and 5 from further processing as they both included the age group 85+. Collapsing Health Regions together is done when there are a few cases of less than 20 respondents but when the problem is wide spread it causes a loss of precision in the data and therefore these scenarios were rejected.

Table 1. A cross Canada selection of Health Regions and their corresponding respondent counts for the selected age groups of 65-74 and 85+.

Table 1 – Number of Respondents of the CCHS in a given Health Region, Age Group and Sex					
Age Group		65-74		85+	
Province	Health Region	Male	Female	Male	Female
Newfoundland and Labrador	Labrador-Grenfell Regional Integrated Health Authority	>=20	>=20	<20	<20
Québec	Région de Bas-Saint-Laurent	>=20	>=20	<20	<20
British Colombia	Vancouver	>=20	>=20	<20	<20
Ontario	City of Toronto Health Unit	>=20	>=20	<20	>=20

As you can see in Table 1, even for the largest HR (City of Toronto Health Unit), the number of respondents, especially for men, does not reach the suitable 20 count minimum for the age group 85+. Therefore we will remove scenario 4 and 5 from further consideration.

The next step is to compare the new person – level weights to those created in production to test the reliability of the results based on a previously accepted standard. This requires looking at the bootstrap weights, where we can compare year-to-year estimates and coefficients of variation (CV). The comparison is based on selected Health Indicators and comparing every combination of indicator by age group, HR and sex with more than 10 respondents. With less than 10 respondents the variance will be inflated and in the end those results cannot be published due to quality issues so they are removed from estimate CV comparisons. The following is a list of Health Indicators that are used for comparisons:

Current Smoker – daily or occasional
Exposure to Second Hand Smoke at home
Asthma
Arthritis
High Blood Pressure
Diabetes
Heavy Drinker
Perceived Health – Poor or Fair
Perceived Health – Good
Perceived Health – Very Good
Perceived Health – Excellent
Perceived Life stress – Quite a lot or Extreme
Have a regular health care provider

Testing if the differences year to year are significant will tell us the impact of creating new age groups causes too great a change in the reliability of the calculations.

Results and Analysis

After attempting all possible age groups and methods, it was concluded that scenario 3, method 1 was the most reliable of the scenarios and methods. With the ten year age gap from scenario 3 it was possible to have more than enough respondents for the two new age groups while still having enough refinement to make production worthwhile. Looking in the tables in Appendix C we can see that there are not many significant differences between the Health Indicators, the age groups, the sex and the Health Region. Table 2 compares the new 2017 estimates to those originally used in production. It shows a very small percentage of significant differences for the new age groups.

Table 2. We can see that there are a few significant differences in the new and 12+ age groups, but it is such a small percentage of the total number of domains that it can be disregarded. Seeing that the age groups of those less than 65 have zero significant differences proves these age groups were not altered in any way.

Table 2 - Total Number of Significant Differences between the new weights and the production weights of a given year for the age groups of Scenario 3

	2017		2016	
Age Group	Number of Significant Differences	Percent Significant (out of 4134 domains)	Number of Significant Differences	Percent Significant (out of 4134 domains)
12-17	0	0.0%	0	0.0%
18-34	0	0.0%	0	0.0%
35-49	0	0.0%	0	0.0%
50-64	0	0.0%	0	0.0%
65-74	2	0.0%	0	0.0%
75+	0	0.0%	16	0.4%
12+	18	0.4%	20	0.5%

The next analysis was to look at the accuracy of the new estimates. Each estimate is derived from multiple respondents so there will be some variability. The new step is to see if this volatility is to the same standard as used in production.

Table 3. The table below compares side by side statistics produced from the coefficients of variation of the selected estimates calculated with the original weights used in production and the new weights. The production statistics should be regarded as the benchmark as they were sufficient to be final production of 2017. Table for 2016 can be found in Appendix D. All weights derived from less than 10 respondents were removed as they are unpublishable and had very large CVs.

Table 3 – Statistics analyzing the coefficient of variation from the weights produced in Scenario 3, Method 1 in comparison to those used in production in 2017		
Statistics	CVs from Scenario 3, Method 1 weights	CVs from Production weights
Mean	20.07	20.15
Standard Deviation	10.42	10.46
CV	51.95	51.89
Variance	108.68	109.35
Standard Error	0.17	0.17
Interquartile Range	14.66	14.61
100% Max	65.56	66.02
75% Q3	27.91	28.04
50% Median	21.21	21.23
25% Q1	13.25	13.43
0% Min	0.00	0.00

The above calculations also produced a Student's t test statistics of 0.02% which indicates that the mean of the CVs of the two years are not significantly different. With the null hypothesis set to the mean of the coefficient of variation of the original weights, we cannot reject the null hypothesis as the two means are not significantly different with a p-value at 99.99%.

The goal of this project is to test if the new age groups provide estimates that are as reliable as in production. We can see that the statistics above are very similar to the statistics used in production meaning that we can rely on our new weights with the same level of accuracy as they do in production.

After agreeing that the new estimates compare well to the production estimates of a given year, we must next compare our 2017 estimates to the new 2016 estimates, like what is normally done in production. Then we can see if the two comparisons are significantly different for our new age groups.

Table 4. With the creation of two new age groups, it is shown that there is no increase in the percentage of significant differences. In production the percent significant is 6.3% for all those 65+, which becomes 6.0% and 5.8% with the new age groups.

Table 4 – A comparison of the total number of significant differences between the 2017 and 2016 weights created with the new age groups and in production				
	Significant differences between the 2017 and 2016 weights created with scenario 3		**Significant differences between the 2017 and 2016 weights created in production**	
Age Group	Number of Significant Differences	Percent Significant (out of 4134 domains)	Number of Significant Differences	Percent Significant (out of 4134 domains)
12-17	134	3.2%	134	3.2%
18-34	218	5.3%	218	5.3%
35-49	249	6.1%	249	6.1%
50-64	325	7.9%	325	7.9%
65-74	248	6.0%	262	6.3%
75+	240	5.8%		

12+	300	7.2%	294	7.1%

The percentages noted above may be smaller but the next step is to look at the coefficients of variation to determine if these decreases are reliable due to the fact that smaller age groups means less respondents which can increase the variance of the estimations.

It can be expected that the coefficient of variation increases as we create new age groups so the extent of the change has to be researched in order to place the proper amount of confidence in the results that are produced.

Table 5. The following table is produced from the coefficients of variation of the weighted estimates. It is only comparing the age group 65+ in Production (on the right) and age groups 65-74 and 75+ (on the left). By the CCHS publishing guidelines, the combination of question response, age group, HR and sex results must have at least 10 respondents. Those with less than 10 respondents were excluded in these calculations to reduce the impact produced by a high variance attributed to a few people.

Table 5 - Statistical measure of the coefficient of variation from the production weights and from the new age groups for all those 65 and older

Statistics	CVs from Production of the age group 65+	CVs from new age groups 65-74 and 75+
Mean	19.7	20.1
Standard Deviation	9.8	10.4
Variance	96.4	108.7
Range	50.8	65.6
Interquartile Range	14.0	14.7
100% Max	50.8	65.6
75% Q3	27.0	27.9
50% Median	20.1	21.2
25% Q1	13.0	13.3
0% Min	0.0	0.0

The increase in variance is to be expected, as eliminating the cases with less than 10 affects the new age groups more than the 65+ age group. The increase in the mean of CVs is very small, if in production the results are produced to satisfaction with these CVs, then we can conclude that the increase in CV does not greatly affect the outcome of the results. Therefore we can state with confidence that the splitting of the age group 65+ does not significantly increase the coefficient of variation, and therefore the conclusions drawn from data in Table 5 are reliable.

Conclusion and Further Research

In conclusion, the comparison between the new weights and the old weights proves that there exists a way of subdividing the age group 65+ while maintaining the high levels of accuracy that are used in production. Of all the scenarios and methods, the age group 65-74 and 75+ using method 1 estimations has the most reliable weights. Scenario 1 and 2 have 5 year age groups, which in certain years is too small.

Scenario 4 and 5 have the age group 85+, which has too few respondents to get accurate weights while maintaining confidentiality. Method 2 was identical to method 1 after a slight revision and method 3 proved ineffective due to a large variance in the population estimates and possibility of negative values.

The age group 65-74 represents people who are likely still living at home, and who rely on less medical services than those in the 75+ age group. This information will be useful for those analyzing the health indicators of these two new age groups, where decisions can be made about these growing populations by grouping like populations.

For future research, the 85+ age group can be analyzed on the provincial level, where the counts are sufficient, to get estimates about the small but growing population. Another solution would be to get more respondents to the 85+ age group, so that this analysis can be done on the health region level. This is the more difficult and costly option but it may provide necessary information as life expectancy increases and more people live well into their 80s and 90s. Perhaps in a few years when the majority of the Baby Boomer generation is in the 85+ age group will it be possible to have a large enough sample of respondents.

Appendix

Appendix A - Proof of Equations used in Method 3

Fact: $(45 \text{ to } 64 \text{ in } 2012) - (45 \text{ to } 59 \text{ in } 2012) = (60 \text{ to } 64 \text{ in } 2012)$

$$\text{Left Side (LS)} = \text{Right Side (RS)}$$

Proof:

$$\begin{aligned}
\text{LS} &= (45 \text{ to } 64 \text{ in } 2012) - (45 \text{ to } 59 \text{ in } 2012) \\
&= (50 \text{ to } 69 \text{ in } 2017) - (50 \text{ to } 64 \text{ in } 2017) + \triangle(2012\text{'s } 45 \text{ to } 64 \text{ cohort}) - \triangle(2012\text{'s } 45 \text{ to } 59 \text{ cohort}) \\
&= (65 \text{ to } 69 \text{ in } 2017) + \triangle(2012\text{'s } 60 \text{ to } 64 \text{ cohort}) \\
&= (60 \text{ to } 64 \text{ in } 2012) \\
&= \text{RS}
\end{aligned}$$

Since LS=RS, the above equality holds and we can rearrange the equation to isolate 2017's 65 to 69 age group. ∎

Fact: $(45 \text{ to } 64 \text{ in } 2012) - (45 \text{ to } 59 \text{ in } 2012) = (60 \text{ to } 64 \text{ in } 2012)$

$$\Leftrightarrow (45 \text{ to } 64 \text{ in } 2012) - [(50 \text{ to } 64 \text{ in } 2017) + \triangle(2012\text{'s } 45 \text{ to } 59 \text{ cohort})] = (65 \text{ to } 69 \text{ in } 2017) + \triangle(2012\text{'s } 60\text{-}64 \text{ cohort})$$

Subtract $\triangle$(2012's 60-64 cohort) from both sides of the equation

$$\Leftrightarrow (45 \text{ to } 64 \text{ in } 2012) - (50 \text{ to } 64 \text{ in } 2017) - \triangle(2012\text{'s } 60\text{-}64 \text{ cohort}) - \triangle(2012\text{'s } 45 \text{ to } 59 \text{ cohort}) = (65 \text{ to } 69 \text{ in } 2017)$$

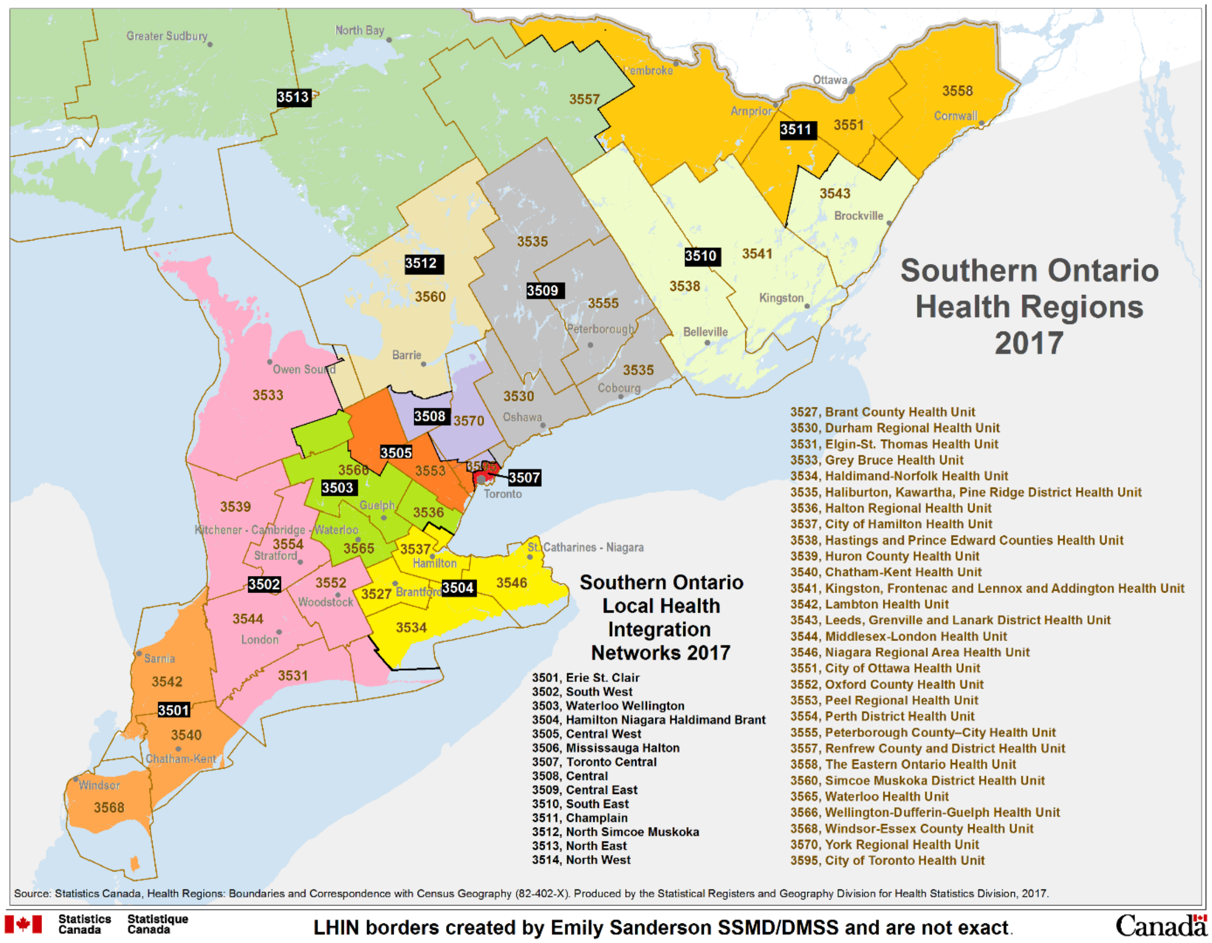
$$\Leftrightarrow \underbrace{(45 \text{ to } 64 \text{ in } 2012) - (50 \text{ to } 64 \text{ in } 2017)}_{\text{Have}} - \underbrace{\triangle(2012\text{'s } 45\text{-}64 \text{ cohort})}_{\text{Need}} = \underbrace{(65 \text{ to } 69 \text{ in } 2017)}_{\text{Want}}$$

The following equations are used in scenario 2 and 3 to calculate the age group 75+.

$$(65+ \text{ in } 2007) + \triangle(2007\text{'s } 65+ \text{ cohort}) = (75+ \text{ in } 2017)$$

This changes are calculated at the provincial level where the counts for each individual age by province, sex and month are available. For example, the estimations for March will be calculated using the provincial population counts of March of 2012 and 2007.

Appendix B



Southern Ontario Health Regions 2017
Southern Ontario Local Health Integration Networks 2017
3501, Erie St. Clair
3502, South West
3503, Waterloo Wellington
3504, Hamilton Niagara Haldimand Brant
3505, Central West
3506, Mississauga Halton
3507, Toronto Central
3508, Central
3509, Central East
3510, South East
3511, Champlain
3512, North Simcoe Muskoka
3513, North East
3514, North West
3527, Brant County Health Unit
3530, Durham Regional Health Unit
3531, Elgin-St. Thomas Health Unit
3533, Grey Bruce Health Unit
3534, Haldimand-Norfolk Health Unit
3535, Haliburton, Kawartha, Pine Ridge District Health Unit
3536, Halton Regional Health Unit
3537, City of Hamilton Health Unit
3538, Hastings and Prince Edward Counties Health Unit
3539, Huron County Health Unit
3540, Chatham-Kent Health Unit
3541, Kingston, Frontenac and Lennox and Addington Health Unit
3542, Lambton Health Unit
3543, Leeds, Grenville and Lanark District Health Unit
3544, Middlesex-London Health Unit
3546, Niagara Regional Area Health Unit
3551, City of Ottawa Health Unit
3552, Oxford County Health Unit
3553, Peel Regional Health Unit
3554, Perth District Health Unit
3555, Peterborough County–City Health Unit
3557, Renfrew County and District Health Unit
3558, The Eastern Ontario Health Unit
3560, Simcoe Muskoka District Health Unit
3565, Waterloo Health Unit
3566, Wellington-Dufferin-Guelph Health Unit
3568, Windsor-Essex County Health Unit
3570, York Regional Health Unit
3595, City of Toronto Health Unit
Greater Sudbury
North Bay
Pembroke
Ottawa
Arnprior
Cornwall
Brockville
Kingston
Belleville
Peterborough
Cobourg
Oshawa
Barrie
Owen Sound
Toronto
Guelph
Kitchener - Cambridge - Waterloo
Stratford
Hamilton
St. Catharines - Niagara
Brantford
Woodstock
London
Sarnia
Chatham-Kent
Windsor
Source: Statistics Canada, Health Regions: Boundaries and Correspondence with Census Geography (82-402-X). Produced by the Statistical Registers and Geography Division for Health Statistics Division, 2017.
Statistics Canada Statistique Canada
LHIN borders created by Emily Sanderson SSMD/DMSS and are not exact.
Canada

Appendix C

Table 6 - Total Number of Significant Differences between Scenario 3, Method 1, 2017 Master and the original 2017 Master by Health Indicator			
Health Indicator	Number of Domains	Number of Significant Differences	Percent Significant
Smokers - Daily or Occasionally	2228	1	0.0%
Exposure to second-hand smoke	2228	1	0.0%
Diabetes	2228	1	0.0%
Asthma	2228	0	0.0%
Arthritis	2228	3	0.1%
High blood pressure	2228	4	0.2%
Have a Regular Health Care Provider	2228	2	0.1%
Perceived health - poor	2228	3	0.1%
Perceived health - good	2228	1	0.0%
Perceived health - very good	2228	0	0.0%
Perceived health - Excellent	2228	1	0.0%
Heavy drinker (of Alcohol)	2228	2	0.1%
Perceived life stress - quite a bit or extreme	2228	1	0.0%

Table 7 -Total Number of Significant Differences between Scenario 3, Method 1, 2017 Master and the original 2017 Master by Health Region			
Table with Health Regions with greater than 0 significant differences			
Health Region	Number of Domains	Number of Significant Differences	Percent Significant
1013	271	1	0.4%
2414	271	3	1.1%
2416	271	4	1.5%
3537	271	1	0.4%
3551	271	1	0.4%
3566	271	1	0.4%
4705	271	3	1.1%
4708	271	3	1.1%
5911	271	1	0.4%
5953	271	1	0.4%
All	271	1	0.4%

Table 8 - Total Number of Significant Differences between Scenario 3, Method 1, 2017 Master and the original 2017 Master by Sex			
Sex	Number of Domains	Number of Significant Differences	Percent Significant
Male	9655	7	0.1%
Female	9655	6	0.1%
All	9655	7	0.1%

The following path contains all of the excel files that have these tables for every scenario and method tested (note method 1 and method 2 are the same)

Appendix D

The following tables are the 2016 version of what was mentioned above.

Table 9 – Statistics analyzing the coefficient of variation of the weights produced in Scenario 3, Method 1 in comparison to those used in production in 2016		
Statistics	CVs from Scenario 3, Method 1 weights	CVs from Production weights
Mean	20.3	20.4
Standard Deviation	10.6	10.6
CV	52.0	51.9
Variance	111.3	111.5
Interquartile Range	15.0	15.1
100% Max	62.6	63.9
75% Q3	28.2	28.4
50% Median	21.1	21.3
25% Q1	13.3	13.4
0% Min	0.0	0.0

References

Deville, J.-C., & Sarndal, C.-E. (1992, June). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association, 87*(418), 376-382. Retrieved October 26, 2009, from Calibration Estimators in Survey Sampling: http://www.jstor.org/stable/2290268

Girard, C. (2009). *A companion to variance estimation written in asymptotically layman terms.* Business Survey Methods Division. Ottawa: Statistics Canada.

Hall, P. (2003). A short prehistory of the bootstrap. *Statistical Science, 18*(2), 158-167.

Statistics Canada. (1992). Survey Methodology. In C. W. J.N.K. Rao, *Some Recent Work on Resampling Methods for Complex Surveys* (2 ed., Vol. 18, pp. 209-217). Ottawa: Statistics Canada.