CPHI

Measuring Health Inequalities

A Toolkit

SAS Macros and Methodology Notes

Canadian Institute for Health Information
Institut canadien d'information sur la santé

Production of this document is made possible by financial contributions from Health Canada and provincial and territorial governments. The views expressed herein do not necessarily represent the views of Health Canada or any provincial or territorial government.

For permission or information, please contact CIHI:

Canadian Institute for Health Information
495 Richmond Road, Suite 600
Ottawa, Ontario K2A 4H6
Phone: 613-241-7860
Fax: 613-241-8120
www.cihi.ca
copyright@cihi.ca

Cette publication est aussi disponible en français sous le titre *Mesurer les inégalités en santé : trousse d’outils — macros SAS et notes méthodologiques*.

Table of contents

This guide provides an overview of 2 SAS macros you can use for analyzing health inequalities. It also includes methodology notes, formulas and considerations for reporting your results.

SAS macros

1. %Calculate_Stratified_Rates

This macro calculates crude and age-standardized rates, stratified by income quintile or geographic location (urban versus rural/remote). Stratified rates are also calculated by sex (male, female and both sexes) and reporting level (national or provincial/territorial). This code may be modified to calculate stratified rates for additional equity stratifiers.

To use the macro

- Create the following 3 input data sets using the format described in the Excel document *Measuring Health Inequalities: A Toolkit — Input File Formats for SAS Macros*:
  1. Indicator cases;
  2. Population denominators; and
  3. A standard population stratified by your age groups (e.g., 5-year age groups).
- Define macro parameters, including the age range for the indicator.
- Run the %Calculate_Stratified_Rates macro (Appendix 1).
- Check that your output is complete and that the log is free of errors and warnings.

2. %Calculate_Inequality_Measures

This macro calculates inequality measures including rate ratio (RR), rate difference (RD), potential rate reduction (PRR) and population impact number (PIN) for health indicator rates stratified by income quintile or urban versus rural/remote. Results are generated by sex (male, female and both sexes) and reporting level (national or provincial/territorial).

To use the macro

- Create an input data set that is formatted as described in the Excel document *Measuring Health Inequalities: A Toolkit — Input File Formats for SAS Macros*. Note that the output data from successfully running %Calculate_Stratified_Rates will be correctly formatted.
- Run the %Calculate_Inequality_Measures macro (Appendix 2).
- Check that your output is complete and that the log is free of errors and warnings.

Methodology notes

1. Age-standardized rates

Indicator rates are age-standardized by the direct method of standardization, using the 2011 Canadian population (from the 2011 Census) as the standard population. Standardization is based on 5-year age groupings.

The **age-standardized rate** is calculated using the following formula:

$$\text{Age-standardized rate} = \frac{\text{Numerator}}{\text{Denominator}} \times \text{Weight} \times \text{Multiplier}$$

where Weight is the weight of the standard population and Multiplier is the value for which you would like to present rates (e.g., per 100,000 population).

The **variance** is calculated using the following formula:

$$\text{Variance (Rate)} = \text{Weight}^2 \times \text{Rate}_{\text{crude}} \times \frac{\left(100{,}000 - \left(\text{Rate}_{\text{crude}}\right)\right)}{\text{Population}}$$

2. Simple inequality measures

Rate ratio is a measure of the relative inequality between subgroups that is calculated by dividing the rate of one subgroup (usually the subgroup with the least desirable rate) by the rate of another subgroup (usually the subgroup with the most desirable rate).

> *Example:* Measuring income-related inequalities, where Q1 is the lowest income quintile and Q5 is the highest income quintile.
>
> Q1 ÷ Q5 = 10 ÷ 5 = 2.00
>
> *Interpretation:* The rate of condition X is 2 times higher for Canadians in the lowest income quintile than for those in the highest income quintile.

The **variance** is calculated using the following formula:

$$\text{Variance}\left(\text{Log}\left(\frac{\text{Rate}_{Q1}}{\text{Rate}_{Q5}}\right)\right) = \frac{\text{Variance (Rate}_{Q1})}{{\text{Rate}_{Q1}}^2} + \frac{\text{Variance (Rate}_{Q5})}{{\text{Rate}_{Q5}}^2}$$

The **RR 95% confidence interval** is given by $\mathbf{e^{(\log\,(Rate\ ratio) \pm 1.96\sqrt{Variance\ (Rate\ ratio)}}}$.

The RR is considered to be statistically significantly different from the null (i.e., RR = 1) if the 95% confidence interval does not include 1.

Rate difference is a measure of the *absolute* inequality between subgroups that is calculated by subtracting the rates of 2 subgroups (usually the highest and the lowest rates).

> *Example:* Q1 − Q5 = 10 − 5 = 5
>
> *Interpretation:* There are 5 more Canadians with condition X in the lowest-income quintile than in the highest-income quintile.

The **variance** is calculated using the following formula:

$$\text{Variance (Rate difference)} = \text{Variance (Rate}_{Q1}) + \text{Variance (Rate}_{Q5})$$

The **RD 95% confidence interval** is given by $\mathbf{Rate\ difference \pm 1.96\sqrt{Variance\ (Rate\ difference)}}$.

The RD is considered to be statistically significantly different from the null (i.e., RD = 0) if the 95% confidence interval does not include 0.

3. Complex inequality measures

A complex inequality measure incorporates data from all population subgroups (e.g., inequality across all income groups); it is a single number indicating the level of inequality.

Potential rate reduction is a *relative* measure of the potential reduction in a health indicator rate that would occur in the hypothetical scenario that each population subgroup experienced the same rate as the subgroup with the most desirable rate. It is also commonly known as the population-attributable fraction or population-attributable risk.[1]

Example: PRR = 15.0%

Interpretation: In a given year, 15% of hospitalizations related to condition X could have been avoided if Canadians in all income quintiles had experienced the same rate of hospitalizations as those in the highest income quintile.

The PRR and its **95% confidence interval** are calculated in the following manner, using the example of a 5-category stratifier (e.g., income quintiles) where category 5 defines the subgroup with the most desirable rate:

$$\text{PRR} = \frac{\sum_{i=1}^{5} P_i \left(\frac{\text{Rate}_i}{\text{Rate}_5} - 1\right)}{1 + \sum_{i=1}^{5} P_i \left(\frac{\text{Rate}_i}{\text{Rate}_5} - 1\right)} \times 100\%$$

where P_i is the proportion of the population in the *i*th category.

The lower and upper confidence intervals are then given as follows:

$$\text{LCI}_{\text{PRR}} = 1 - \frac{1}{P_5 + e^{\left(\log\left(\text{Rate ratio}_{1:4,\,5}\right) - 1.96\sqrt{\text{Variance}\left(\text{Rate ratio}_{1:4,\,5}\right)}\right)}}$$

$$\text{UCI}_{\text{PRR}} = 1 - \frac{1}{P_5 + e^{\left(\log\left(\text{Rate ratio}_{1:4,\,5}\right) + 1.96\sqrt{\text{Variance}\left(\text{Rate ratio}_{1:4,\,5}\right)}\right)}}$$

where Rate $\text{ratio}_{1:4,\,5}$ is the ratio of the sum of the population proportion (P_i) multiplied by the age-standardized rate (Rate_i) in the first category relative to the rate in the reference category, calculated as

$$\frac{\sum_{i=1}^{4} P_i \text{Rate}_i}{\text{Rate}_5}$$

with variance

$$\frac{\sum_{i=1}^{4} P_i^2 \times \text{Var}\left(\text{Rate}_i\right)}{\left(\sum_{i=1}^{4} P_i \text{Rate}_i\right)^2} + \frac{\text{Variance}\left(\text{Rate}_5\right)}{\text{Rate}_5^{\,2}}$$

The PRR is considered to be statistically significantly different from the null (i.e., PRR = 0) if the 95% confidence interval does not include 0.

Population impact number is an *absolute* measure of the potential reduction in the number of cases for a health indicator that would occur in the hypothetical scenario that each population subgroup experienced the same rate as the subgroup with the most desirable rate. It captures the gradient of inequality across multiple categories, such as income quintiles.

> *Example:* PIN = 7,300
>
> *Interpretation:* In a given year, 7,300 hospitalizations related to condition X could have been avoided if Canadians in all income levels had experienced the same rate of hospitalizations as those in the highest income level.

The PIN is related to the PRR in the following manner:

$$\text{Population impact number} = \text{Overall indicator standardized rate} \times N_{\text{total population}} \times \text{PRR}$$

Considerations for reporting complex inequality measures

- The PRR is ideally suited for scenarios where lower rates of an outcome are desirable and the RR is greater than 1. For indicators where higher rates are desirable and the RR is less than 1, consider calculating a potential rate improvement (PRI, also known as the prevented fraction).
- Consider reporting both the PRR and the PIN as 0 if the PRR is negative and statistically different from 0 (i.e., upper and lower confidence limits are negative).
- Consider reporting the PIN as an approximate number (e.g., rounded to the nearest 100).

Appendix 1: SAS macro — %Calculate_Stratified_Rates

Macro name

Calculate_Stratified_Rates

Purpose

To calculate crude and age-standardized rates, stratified by income quintile or geographic location (urban versus rural/remote). Income and geographic location (urban versus rural/remote) are 2 commonly used equity stratifiers for measuring health inequalities. Stratified rates are also calculated by sex (male, female and both sexes) and reporting level (national or provincial/territorial). This code may be modified to calculate stratified rates for additional equity stratifiers.

Input files

In order for this macro to work, you will need 3 input data sets:

1. Indicator cases — by equity stratifier, reporting level, sex and age group
2. Population denominators — by equity stratifier, reporting level, sex and age group
3. A standard population — we use the 2011 Canadian standard population by 5-year age groups

All input data sets should be formatted as specified in the Excel document *Measuring Health Inequalities: A Toolkit — Input File Formats for SAS Macros*.

Output file

&indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates — Crude and age-standardized rates for the chosen equity stratifier (income or geographic location) and the reporting level.

Parameters (all required)

Prior to running the macro, you must define the &indata_directory and &outdata_directory, as well as the parameters listed below.

- *indicator* — The name of the health indicator you are measuring; this can be any name you choose
- *yr* — Year
- *reporting_level* — The geographic reporting level (must be specified as province_code)
- *equity_stratifier* — The variable for which you would like to calculate stratified rates; there are currently 2 options: QAIPPE for income quintile or URR for geographic location (urban and rural/remote)
- *infile* — The name of the input data set of indicator cases
- *stdpop* — The name of the standard population data set
- *popfile* — The name of the population data set corresponding to income or geographic location (i.e., population estimates by income quintile or population estimates by urban and rural/remote geographic location)
- *popvar* — The population counts from popfile (must be specified as province_pop)
- *outfile* — The name of the output data set
- *AGEGPL* — The lower age limit of the health indicator; must be in format 01, 02 . . . 14, 15 as specified in the age variable (e.g., if the lower age limit for your health indicator is 0 years, this value would be 01).
- *AGEGPU* — The upper age limit of the health indicator; must be in format 01, 02 . . .14, 15 as specified in the age variable (e.g., if the upper age limit for your health indicator is 74 years, this value would be 15).
- *MULTIPLIER* — The value for which you would like to present rates (e.g., MULTIPLIER = 100,000 for rates expressed per 100,000)
- *ROUNDUNIT* — The rounding unit (e.g., to round to 1 decimal place, specify ROUNDUNIT = .01)

Macro invocation

%MACRO Calculate_Stratified_Rates (indicator =, yr =, reporting_level =, equity_stratifier =, infile =, stdpop =, popfile =, popvar =, outfile =, AGEGPL =, AGEGPU =, MULTIPLIER =, ROUNDUNIT =);

Steps

1. Prepare the indicator cases by creating aggregated counts for reporting level, age group, sex and equity stratifier.
2. Prepare the standard population file used for age standardization.
3. Prepare the population denominators for the age range of interest.
4. Link the 3 data sets listed above at the provincial/territorial level by age group, sex and equity stratifier.
5. Calculate stratified crude rates and use the following steps to calculate stratified age standardized rates:
    a) Calculate age-specific rates for each age group.
    b) Multiply the age-specific rates of the population under study by the number of persons in each age group of the standard population to get the expected value for each age group.
    c) Add the number of expected values from all age groups. Finally, to get the age-adjusted rates, divide the total expected value by the standard population.
    d) Calculate variance and use this to calculate confidence intervals.
6. Calculate overall crude and age-standardized rates. Note that overall rates refer to all categories combined within the equity stratifier (e.g., for the income stratifier, this refers to the overall rate for quintiles 1 through 5). For this reason, any cases that are not assigned to an income category (e.g., for missing postal code) will be excluded from the overall rate.
7. Output the data table.

Macro code

```
/*=========================================================================*/
/* - Define input library
/*=========================================================================*/
libname data "&indata_directory.";
libname out "&outdata_directory.";

/*=========================================================================*/
/* - Define %include statements
/*=========================================================================*/


%MACRO Calculate_Stratified_Rates(indicator= , yr= , reporting_level= ,
equity_stratifier= , infile=, stdpop=,    popfile=,   popvar=,
outfile=,AGEGPL=, AGEGPU=, MULTIPLIER=,   ROUNDUNIT=);

/*****************************************************************/
/* STEP 1: Prepare data set of individual-level indicator cases   */
/*****************************************************************/

/*****************************************************************/
/*          1A. Create count variable for the indicator data set and output
required year(s) and 5-year age groups
/*****************************************************************/
data &indicator.&yr._&AGEGPL._&AGEGPU.;
      set &infile.;
      * Create a count variable for the indicator;
      &indicator.=1;
      * Output required year and age groups only;
      if  year eq &yr. and ("&AGEGPL."<=AGE_GROUP_CODE<="&AGEGPU.") then
output;
run;

/*****************************************************************/
/*          1B. Create aggregated indicator data sets - national and by
province/territory       */
/*    ==> Prepare aggregated counts from indicator data set (infile) by age
group, sex (male/female/overall) and equity stratifier.
/*    Do this for Canada and by reporting level (&reporting_level.)
/*****************************************************************/
* Sum indicator cases by sex, age group and equity stratifier for all of
Canada;
proc sql;
      create table &indicator._canada as
      select  sex, AGE_GROUP_CODE, &equity_stratifier., SUM(&indicator.) as
cases
      from &indicator.&yr._&AGEGPL._&AGEGPU.
      group by  sex, AGE_GROUP_CODE, &equity_stratifier.
      OUTER UNION CORR
      select "3" AS sex, AGE_GROUP_CODE, &equity_stratifier.,
SUM(&indicator.) as cases
      from &indicator.&yr._&AGEGPL._&AGEGPU.
      group by  AGE_GROUP_CODE, &equity_stratifier.
      ORDER BY SEX, AGE_GROUP_CODE, &equity_stratifier.;
QUIT;
```

```
/* Define Canada as '99' */
data &indicator._canada;
      set &indicator._canada;
      &reporting_level.="99";
run;

* Sum indicator cases by sex, age group and equity stratifier by reporting
level from &reporting_level.;
proc sql;
      create table &indicator._&reporting_level. as
      select  &reporting_level., sex, AGE_GROUP_CODE, &equity_stratifier.,
SUM(&indicator.) as cases
      from &indicator.&yr._&AGEGPL._&AGEGPU.
      group by  &reporting_level., sex, AGE_GROUP_CODE, &equity_stratifier.
      OUTER UNION CORR
      select &reporting_level., "3" AS sex, AGE_GROUP_CODE,
&equity_stratifier., SUM(&indicator.) as cases
     from &indicator.&yr._&AGEGPL._&AGEGPU.
      group by  &reporting_level., AGE_GROUP_CODE, &equity_stratifier.
      ORDER BY &reporting_level., SEX, AGE_GROUP_CODE, &equity_stratifier.;
QUIT;

data &indicator._&reporting_level.;
      set &indicator._&reporting_level. &indicator._canada;
run;

/****************************************************************/
/*          STEP 2. Prepare standard population file            */
/****************************************************************/

/****************************************************************/
/*          2A. Sum the standard population by indicator age groups within
age range of the indicator and keep the total population
/*          ==>  Multiply each of the age-specific rates by the proportion of
the 2011 population belonging to the particular age group (called the
standard population weight)
/****************************************************************/

proc sql;
  create table stdpop_agegroup as
  select AGE_GROUP_CODE, standard_pop_cnt as stdpop, sum(standard_pop_cnt) as
agegroup_stdpop
  from &stdpop.(where=("&AGEGPL."<=AGE_GROUP_CODE<="&AGEGPU."));
quit;
proc sort data=stdpop_agegroup; by AGE_GROUP_CODE; run;

proc sql;
  create table stdpop_population as
  select AGE_GROUP_CODE, sum(standard_pop_cnt) as total_stdpop
  from &stdpop.(where=("&AGEGPL."<=AGE_GROUP_CODE<="&AGEGPU."));
quit;
proc sort data=stdpop_population; by AGE_GROUP_CODE; run;

data stdpop_new;
      merge stdpop_agegroup (in=a) stdpop_population(in=b);
      by AGE_GROUP_CODE;
```

```
      if a=1  then do;
            weight= stdpop/agegroup_stdpop;
            output;
      end;
      drop total_stdpop;
run;

/*****************************************************************/
/*          STEP 3. Prepare population data              */
/*****************************************************************/

/*****************************************************************/
/*          3A. Prepare population denominators for the specific age range of
indicator */
/*****************************************************************/
data new_popfile;
      set &popfile.;
      if  population_year eq &yr. and
("&AGEGPL."<=AGE_GROUP_CODE<="&AGEGPU.") then output;
run;

/*****************************************************************/
/*          STEP 4. Join data sets from steps 1 to 3
*/
/*****************************************************************/
/*****************************************************************/
/*          4A. Join standard population and population file
/*****************************************************************/
proc sql;
  create table new_popfile2 as
  select a.*, b.stdpop, b.weight
  from  new_popfile as a
  right join stdpop_new as b
  on    a.AGE_GROUP_CODE=b.AGE_GROUP_CODE
  order by &reporting_level., a.sex, b.AGE_GROUP_CODE, a.&equity_stratifier.
;
quit;


/*****************************************************************/
/*          4B. Add in aggregated numerator cases and roll up by equity
stratifier, reporting_level, sex and age
/*****************************************************************/
proc sql;
        create table count_pop_&equity_stratifier. as
        select a.cases as count, b.*
        from   &indicator._&reporting_level. as a
        right join new_popfile2 as b
on a.AGE_GROUP_CODE=b.AGE_GROUP_CODE and a.sex=b.sex and
a.&equity_stratifier.= b.&equity_stratifier. and
a.&reporting_level.=b.&reporting_level.
        order by &reporting_level., &equity_stratifier., sex, AGE_GROUP_CODE;
quit;
```

```
/*******************************************************************/
/*          4C. Create new population file rolling up all income quintiles or
geographic locations (urban and rural/remote) to be used in calculating
overall age-standardized rates
/*******************************************************************/
proc sql;
        create table new_popfile3 as
        select population_year, province_code, sex, age_group_code,
sum(&popvar.) as &popvar., stdpop, weight
        from   new_popfile2
        group by population_year, province_code, sex, age_group_code, stdpop,
weight;
quit;

/*******************************************************************/
/*          4D. Add in aggregated numerator cases by reporting level, sex and
age only - will use this for calculating overall rates (e.g., rates for all
income quintiles combined or urban and rural/remote combined)
/*******************************************************************/

proc sql;
      create table &indicator._&reporting_level._all
      as select &reporting_level., sex, AGE_GROUP_CODE, sum(cases) as cases
      from &indicator._&reporting_level.
      group by &reporting_level., sex, AGE_GROUP_CODE;
quit;

proc sql;
        create table count_pop as
        select a.cases as count, b.*
        from   &indicator._&reporting_level._all as a
        right join new_popfile3 as b
        on a.AGE_GROUP_CODE=b.AGE_GROUP_CODE and a.sex=b.sex and
a.&reporting_level.=b.&reporting_level.
        order by &reporting_level., sex, AGE_GROUP_CODE;
quit;


/*******************************************************************/
/*    STEP 5. Calculate crude and age-standardized rates stratified by
&equity_stratifier
/*******************************************************************/

/*******************************************************************/
/*    5A. Calculate crude, expected, variance and standardized numerator
count for each row in data set
/*          ==> Variance calculation is based on a binomial distribution
/*          ==> Standardized numerator count (std_num) is required in
Calculate_Inequality_Measures macro */
/*******************************************************************/
%if &popvar. ne 0  %then %do;
 data &indicator._rates_&YR._&AGEGPL._&AGEGPU.;
      set count_pop_&equity_stratifier.;
      if count=. then count=0;
      crude=(count/&popvar.)*&MULTIPLIER.;
      expect=crude*weight;
      variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
```

```
      std_num=count/&popvar.*stdpop;
 run;
%end;

%if &popvar. = 0  %then %do;
 data &indicator._rates_&YR._&AGEGPL._&AGEGPU.;
      set count_pop_&equity_stratifier.;
        if count=. then count=0;
        crude=0;
        expect=crude*weight;
        variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
        std_num=count/&popvar.*stdpop;
 run;
%end;

/*****************************************************************/
/*    5B. Calculate age-standardized rates (ASR) by equity stratifier
/*          ==>  Sum counts, total population, crude rate, standard
numerator, expected and variance by
/*          ==>  geographic reporting level, sex and equity stratifier
/*****************************************************************/
proc sql;
  create table &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.  as
      select &reporting_level., sex, &equity_stratifier.,
      sum(count) as count, sum(&popvar.) as totalpop,
      sum(count)/sum(&popvar.)*&MULTIPLIER. as crude, sum(std_num) as
std_numer,
      sum(expect) as sum_exp, sum(variance) as variance
    from  &indicator._rates_&YR._&AGEGPL._&AGEGPU.
  group by &reporting_level., sex, &equity_stratifier.;
quit;

/*****************************************************************/
/*    5C. Calculate confidence intervals for age-standardized rates by equity
stratifier
/*****************************************************************/
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
      set &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
      lci=sum_exp-1.96*sqrt(variance);
      uci=sum_exp+1.96*sqrt(variance);

      if lci<0 then do;
              p=sum_exp/&MULTIPLIER.;
              if p=0 then p=1/(&MULTIPLIER.*&MULTIPLIER.);
              logitp=log(p/(1-p));
              var=1/(p*p*(1-p)*(1-p))*(variance/(&MULTIPLIER.*&MULTIPLIER.));
              A=logitp-1.96*sqrt(var);
              B=logitp+1.96*sqrt(var);
              lci=&MULTIPLIER./(1+exp(-A));
              uci=&MULTIPLIER./(1+exp(-B));
              drop p logitp var A B;
end;
%if &ROUNDUNIT. = 1 %then
            format crude rstd rlci ruci 8.;
%else
format asr asr_lci asr_uci 8&ROUNDUNIT.;;
rcrude= round(crude,&ROUNDUNIT.);
```

```
asr = round(sum_exp,&ROUNDUNIT.);
asr_lci = round(lci,&ROUNDUNIT.);
asr_uci = round(uci,&ROUNDUNIT.);
run;

/*****************************************************************/
/*    STEP 6.     Overall rates                                  */
/*****************************************************************/

/*****************************************************************/
/*    6A. Calculate crude, expected, variance and standardized numerator
count for each row in data set
/*          ==>  Variance calculation is based on a binomial distribution
/*          ==>  Standardized numerator count (std_num) is required in
Calculate_Inequality_Measures macro */
/*****************************************************************/
%if &popvar. ne 0  %then %do;
 data &indicator._rates_&YR._&AGEGPL._&AGEGPU._ALL;
      set count_pop;
      if count=. then count=0;
      crude=(count/&popvar.)*&MULTIPLIER.;
      expect=crude*weight;
      variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
      std_num=count/&popvar.*stdpop;
 run;
%end;

%if &popvar. = 0  %then %do;
 data &indicator._rates_&YR._&AGEGPL._&AGEGPU._ALL;
        set count_pop;
        if count=. then count=0;
        crude=0;
        expect=crude*weight;
        variance=((weight*weight)*crude*(&MULTIPLIER.-crude))/&popvar.;
        std_num=count/&popvar.*stdpop;
 run;
%end;

/*****************************************************************/
/*    6B. Calculate age-standardized rates (ASR) by equity stratifier
/*          ==> Sum counts, total population, crude rate, standard numerator,
expected and variance by
/*          ==>  geographic reporting level, sex and equity stratifier
/*****************************************************************/
proc sql;
  create table &indicator._ASR_&YR._&AGEGPL._&AGEGPU. as
      select &reporting_level., sex,
      sum(count) as count, sum(&popvar.) as totalpop,
      sum(count)/sum(&popvar.)*&MULTIPLIER. as crude, sum(std_num) as
std_numer,
      sum(expect) as sum_exp, sum(variance) as variance
    from  &indicator._rates_&YR._&AGEGPL._&AGEGPU._ALL
  group by &reporting_level., sex;
quit;
```

```
/*****************************************************************/
/*    6C. Calculate confidence intervals for overall age-standardized rates
/*****************************************************************/
data  &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
      set &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
      lci=sum_exp-1.96*sqrt(variance);
      uci=sum_exp+1.96*sqrt(variance);

      if lci<0 then do;
              p=sum_exp/&MULTIPLIER.;
              if p=0 then p=1/(&MULTIPLIER.*&MULTIPLIER.);
              logitp=log(p/(1-p));
              var=1/(p*p*(1-p)*(1-p))*(variance/(&MULTIPLIER.*&MULTIPLIER.));
              A=logitp-1.96*sqrt(var);
              B=logitp+1.96*sqrt(var);
              lci=&MULTIPLIER./(1+exp(-A));
              uci=&MULTIPLIER./(1+exp(-B));
              drop p logitp var A B;
end;
%if &ROUNDUNIT. = 1 %then
            format crude rstd rlci ruci 8.;
%else
format asr asr_lci asr_uci 8&ROUNDUNIT.;;
rcrude= round(crude,&ROUNDUNIT.);
asr = round(sum_exp,&ROUNDUNIT.);
asr_lci = round(lci,&ROUNDUNIT.);
asr_uci = round(uci,&ROUNDUNIT.);
run;

/*****************************************************************/
/*     STEP 7. Output data set
       Note: This data set can be used as the input data set for the
Calculate_Inequality_Measures macro to calculate inequality summary measures
by equity stratifier.
*/
/*****************************************************************/
* ASR by equity stratifier;
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
      retain fiscal_year &reporting_level. sex &equity_stratifier count
      totalpop rcrude variance asr asr_lci asr_uci std_numer;
      set &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.;
      fiscal_year=&yr.;
      keep fiscal_year &reporting_level. sex &equity_stratifier count
totalpop rcrude variance asr asr_lci asr_uci std_numer;
run;

* ASR overall;
data &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
      retain fiscal_year &reporting_level. sex &equity_stratifier count
      totalpop rcrude variance asr asr_lci asr_uci std_numer;
      set &indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
      fiscal_year=&yr.;
      &equity_stratifier ="0";
      keep fiscal_year &reporting_level. sex &equity_stratifier count
totalpop rcrude variance asr asr_lci asr_uci std_numer;
run;
```

```
* Combine ASR overall and by equity stratifier;
data &indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates;
      set &indicator._ASR_&YR._&AGEGPL._&AGEGPU._&equity_stratifier.
&indicator._ASR_&YR._&AGEGPL._&AGEGPU.;
run;

proc sort data=&indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates;
by &reporting_level. sex &equity_stratifier.; run;

      data out.&outfile.;
            set &indicator._&YR._&AGEGPL._&AGEGPU._&equity_stratifier._rates;
      run;

%MEND Calculate_Stratified_Rates;
```

Appendix 2: SAS macro — %Calculate_Inequality_Measures

Macro name

Calculate_Inequality_Measures

Purpose

To calculate inequality measures including rate ratio (RR), rate difference (RD), potential rate reduction (PRR), and population impact number (PIN) for health indicator rates stratified by income quintile or urban versus rural/remote. Income and geographic location (urban versus rural/remote) are 2 commonly used equity stratifiers for measuring health inequalities. This code calculates income-related and geography-related inequalities for age-standardized rates, by sex and by reporting level (national or provincial/territorial).

Input file format

The input data set should be formatted as specified in the Excel document *Measuring Health Inequalities: A Toolkit — Input File Formats for SAS Macros*. Note that the output data from successfully running %Calculate_Stratified_Rates will be correctly formatted.

Output file created

2 output files will be created after running this macro program:

rd_rr_&indicator.&yr_by_&equity_stratifier. — RR and RD results based on stratifier and reporting level

prr_pin_&indicator.&yr_by_&equity_stratifier. — PRR and PIN results based on stratifier and reporting level

Parameters (all required)

- *data* — The name of the data table being input into the macro
- *indicator* — The name of the health indicator you are measuring; this can be any name you choose
- *yr* — Year
- *equity_stratifier* — The variable for which you would like to calculate inequality measures; there are currently 2 options: QAIPPE for income quintile or URR for geographic location (urban and rural/remote)
- *reporting_level* — The geographic reporting level (must be specified as province_code)
- *sex* — The variable that denotes the sex breakdown of the indicator; must have this even if reporting for both sexes (use a dummy in its place)
- *outlib* — The library name for the output location
- *multiplier* — The value for which you would like to present rates (e.g., multiplier = 100,000 for rates expressed per 100,000)

Macro invocation

%MACRO Calculate_Inequality_Measures (data =, indicator =, yr =, equity_stratifier =, reporting_level =, sex =, outlib =, multiplier =);

Steps

1. Create subsets of the input data for comparing the subgroups with the most desirable and least desirable rates.
2. Calculate RR and RD.
3. Calculate PRR and PIN.
4. Output data tables with all summary measures.

Macro code

```
%MACRO Calculate_Inequality_Measures(data=, indicator=, yr=,
equity_stratifier=, reporting_level=, sex=, outlib=, multiplier=);

/*****************************************************************/
/*             STEP 1: Subset data          */
/*****************************************************************/

%if &equity_stratifier. = QAIPPE %then %do;
/*****************************************************************/
/*** 1A. Subset data by income stratifier to keep Q1 and Q5 ***/

/*****************************************************************/

data &indicator.&yr._lo;
set &data.;
keep &reporting_level. &sex. QAIPPE asr variance asr_uci asr_lci;
if QAIPPE = 1;
run;

data &indicator.&yr._hi;
set &data.;
keep &reporting_level. &sex. QAIPPE asr variance asr_uci asr_lci;
if QAIPPE = 5;
run;
%end;

%if &equity_stratifier. = URR %then %do;
/*****************************************************************/
/*** 1B. Subset data by geography stratifier to keep urban versus
rural_remote ***/
/*****************************************************************/

data &indicator.&yr._hi;
set &data.;
keep &reporting_level. &sex. asr variance asr_uci asr_lci;
if urr = 1;
run;

data &indicator.&yr._lo;
set &data.;
keep &reporting_level. &sex. asr variance asr_uci asr_lci;
if urr = 2;
run;
%end;

/*****************************************************************/
/*  STEP 2: Calculate RR and RD based on stratified data */
*merge back together into flat file to calculate RR, RD*;
/*****************************************************************/
```

```
/******************************************************************/
*2A. Output table with RR + RD*;
/******************************************************************/

proc sql;
create table &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier. as
select &yr. as fiscal_year, a.&reporting_level., a.&sex., a.asr as
lo_std_rate, a.asr_lci as lo_lcl, a.asr_uci as lo_ucl, b.asr as hi_std_rate,
b.asr_lci as hi_lcl, b.asr_uci as hi_ucl, a.variance as lo_var, b.variance as
hi_var,
(a.asr/b.asr) as rr, ((lo_var/(lo_std_rate**2))+(hi_var/(hi_std_rate**2))) as
var_logdrr,
exp(log(calculated rr)-(1.96*sqrt(calculated var_logdrr))) as lcl_rr,
exp(log(calculated rr)+(1.96*sqrt(calculated var_logdrr))) as ucl_rr,
(a.asr - b.asr) as rd, (lo_var+hi_var) as var_rd,
(calculated rd-(1.96*sqrt(calculated var_rd))) as lcl_rd,
(calculated rd+(1.96*sqrt(calculated var_rd))) as ucl_rd
from &indicator.&yr._lo as a inner join &indicator.&yr._hi as b
on a.&reporting_level. = b.&reporting_level. and a.&sex. = b.&sex.;

/******************************************************************/
/* STEP 3: Calculate PRR and PIN based on stratified data (income stratifier
only) */
/******************************************************************/

/******************************************************************/
*3A. Calculate PRR*;
*Use study pop as the population proportion*;
/******************************************************************/

%if &equity_stratifier. = QAIPPE %then %do;
proc sql;
*create column of quintile 5 std rate and proportions*;
create table prr_&indicator.&yr._a as
select &reporting_level., &sex., QAIPPE, totalpop, asr, variance,
(totalpop/sum(totalpop)) as pop_portion
from &data.
where QAIPPE ne '0'
group by &reporting_level., &sex.;

create table prr_&indicator.&yr._b as
select a.&reporting_level., a.&sex., a.QAIPPE, a.totalpop, a.asr, a.variance,
a.pop_portion,
b.asr as tot_std_hi, a.asr/tot_std_hi as rate_ratio_hi
from prr_&indicator.&yr._a as a inner join &indicator.&yr._hi as b
on a.&reporting_level. = b.&reporting_level. and a.&sex. = b.&sex.;

/******************************************************************/
*3B. Calculate PRR*;
/******************************************************************/

create table prr_&indicator.&yr. as
select &reporting_level., &sex., sum(pop_portion*(rate_ratio_hi-
1))/(1+sum(pop_portion*(rate_ratio_hi-1))) as prr
from prr_&indicator.&yr._b
group &reporting_level., &sex.;
```

```
*Calculate CIs for PRR*;
create table rate_low_&indicator.&yr. as
select &reporting_level., &sex., sum((pop_portion**2)*variance) as
variance_low, sum(pop_portion*asr) as sum_low
from prr_&indicator.&yr._b
where QAIPPE ne '5'
group &reporting_level., &sex.;

create table rate_hi_&indicator.&yr. as
select &reporting_level., &sex., variance as variance_hi, asr as rate_hi,
pop_portion as pop_portion_hi
from prr_&indicator.&yr._b
where QAIPPE = '5';

create table prr_ci_&indicator.&yr. as
select a.&reporting_level., a.&sex., prr,
(sum_low/rate_hi) as rate_low,
((variance_low/sum_low**2)+(variance_hi/rate_hi**2)) as varlog_low,
exp(log(calculated rate_low)-1.96*sqrt(calculated varlog_low)) as lcl_low,
exp(log(calculated rate_low)+1.96*sqrt(calculated varlog_low)) as ucl_low,
1-(1/(pop_portion_hi+calculated lcl_low)) as lcl_prr,
1-(1/(pop_portion_hi+calculated ucl_low)) as ucl_prr
from prr_&indicator.&yr. as a, rate_low_&indicator.&yr. as b,
rate_hi_&indicator.&yr. as c
where a.&reporting_level.=b.&reporting_level.=c.&reporting_level. and
a.&sex.=b.&sex.=c.&sex.;

/*****************************************************************/
/*STEP 3: Calculate PRR and PIN based on stratified data (income stratifier
only) */
/*****************************************************************/

create table prr_pin_&indicator.&yr._by_&equity_stratifier. as
select &yr. as fiscal_year, a.&reporting_level., a.&sex.,
%if &equity_stratifier. = QAIPPE %then %do;
      qaippe,
%end;
%if &equity_stratifier. = URR %then %do;
      urr,
%end;
prr, lcl_prr, ucl_prr, std_numer*prr as pin
from &data. as a left join prr_ci_&indicator.&yr. as b
on a.&reporting_level.=b.&reporting_level. and a.&sex.=b.&sex.;

proc sort data=prr_pin_&indicator.&yr._by_&equity_stratifier.
out=&outlib..prr_pin_&indicator.&yr._by_&equity_stratifier. nodup;
      by &reporting_level. &sex.;
run;
%end;

/*****************************************************************/
/* STEP 4: Output final data as a table with all summary measures */
/*****************************************************************/

%if &equity_stratifier. = QAIPPE %then %do;
proc sql;
```

```
create table &indicator._inequality_summary as
select a.fiscal_year, b.&reporting_level. as jurisdiction, c.qaippe, a.&sex.,
b.count, b.rcrude as crude_rate, b.asr,
b.asr_lci, b.asr_uci , a.rr, a.lcl_rr, a.ucl_rr, a.rd, a.lcl_rd, a.ucl_rd,
c.prr, c.lcl_prr, c.ucl_prr, c.pin
from &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier. as a, &data. as
b, &outlib..prr_pin_&indicator.&yr._by_&equity_stratifier. as c
where a.&reporting_level.=b.&reporting_level.=c.&reporting_level. and
a.&sex.=b.&sex.=c.&sex. and b.qaippe = c.qaippe
group by c.qaippe;
%end;

%if &equity_stratifier. = URR %then %do;
proc sql;
create table &indicator._inequality_summary as
select a.fiscal_year, b.&reporting_level. as jurisdiction, b.urr, a.&sex.,
b.count, b.rcrude as crude_rate, b.asr,
b.asr_lci, b.asr_uci , a.rr, a.lcl_rr, a.ucl_rr, a.rd, a.lcl_rd, a.ucl_rd
from &outlib..rd_rr_&indicator._&yr._by_&equity_stratifier. as a, &data. as b
where a.&reporting_level.=b.&reporting_level. and a.&sex.=b.&sex.
group by b.urr;
%end;

%let stratifier2 = %substr(&equity_stratifier.,1,3);

data &outlib..&indicator._&stratifier2._inequality_summary;
      retain fiscal_year jurisdiction Stratifier sex count crude_rate asr
asr_lci asr_uci rr lcl_rr ucl_rr rd lcl_rd ucl_rd
      %if &equity_stratifier. = QAIPPE %then %do;
      prr lcl_prr ucl_prr pin
      %end;
      ;
      length stratifier $12.;
      set &indicator._inequality_summary;
      %if &equity_stratifier. = QAIPPE %then %do;
            if qaippe = '0' then Stratifier = '0 (overall)';
            if qaippe = '1' then Stratifier = '1 (lowest)';
            if qaippe = '2' then Stratifier = '2';
            if qaippe = '3' then Stratifier = '3';
            if qaippe = '4' then Stratifier = '4';
            if qaippe = '5' then Stratifier = '5 (highest)';
            label Stratifier = 'Income Quintile';
      %end;
      %if &equity_stratifier. = URR %then %do;
            if urr eq 0 then Stratifier = '0 (overall)';
            if urr eq 1 then Stratifier = 'Urban';
            if urr eq 2 then Stratifier = 'Rural/Remote';
      %end;
      label fiscal_year = 'Fiscal Year'
            jurisdiction = 'Jurisdiction'
            sex      = 'Sex'
            count    = 'Number of Indicator Cases'
            crude_rate  = "Crude Rate per &MULTIPLIER.  population"
            asr      = "Age-standardized rate per &MULTIPLIER. population"
            asr_lci = 'Lower confidence limit of the age-standardized rate'
            asr_uci = 'Upper confidence limit of the age-standardized rate'
            rr       = 'Rate Ratio'
```

```
            lcl_rr  = 'Lower confidence limit of the rate ratio'
            ucl_rr  = 'Upper confidence limit of the rate ratio'
            rd      = "Rate difference per &MULTIPLIER.  population"
            lcl_rd  = 'Lower confidence limit of the rate difference'
            ucl_rd  = 'Upper confidence limit of the rate difference'
      %if &equity_stratifier. = QAIPPE %then %do;
            prr     = 'Potential Reduction Rate'
            lcl_prr = 'Lower confidence limit of the Potential Rate
      Reduction'
            ucl_prr = 'Upper confidence limit of the Potential Rate
      Reduction'
            pin     = "Population Impact Number per &MULTIPLIER. population";
      %end;
      keep fiscal_year jurisdiction Stratifier sex count crude_rate asr
      asr_lci asr_uci rr lcl_rr ucl_rr rd lcl_rd ucl_rd
      %if &equity_stratifier. = QAIPPE %then %do;
      prr lcl_prr ucl_prr pin
      %end;
      ;
      %if &equity_stratifier. = URR %then %do;
      drop urr;
      %end;
run;

proc sort data=&outlib..&indicator._&stratifier2._inequality_summary;
      by fiscal_year jurisdiction sex Stratifier;
run;
%mend;
```

Reference

1. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *American Journal of Public Health*. January 1998.

help@cihi.ca

CIHI Ottawa
495 Richmond Road
Suite 600
Ottawa, Ont.
K2A 4H6
613-241-7860

CIHI Toronto
4110 Yonge Street
Suite 300
Toronto, Ont.
M2P 2B7
416-481-2002

CIHI Victoria
880 Douglas Street
Suite 600
Victoria, B.C.
V8W 2B7
250-220-4100

CIHI Montréal
1010 Sherbrooke Street West
Suite 602
Montréal, Que.
H3A 2R7
514-842-2226

cihi.ca

18433-0918