

A Macro for the Calculation of Age, Gender and Race Adjusted Incidence and Mortality Rates

Padmini Sekar, Department of Environmental Health, University of Cincinnati.

Matthew Flaherty, Department of Neurology, University of Cincinnati

Daniel Woo, Department of Neurology, University of Cincinnati

Ralph Buncher, Department of Environmental Health, University of Cincinnati,
Cincinnati, Ohio

ABSTRACT:

Two aims of population based studies are to report mortality rates and incidence rates of diseases. This macro was developed by expanding a methodology of an earlier report to calculate age and sex adjusted rates. It now allows the calculation of age, race and sex adjusted rates.

The macro was originally developed to calculate incidence rate of strokes in the five county regions in the Greater Cincinnati area. Since there was no established methodology to calculate rates in biracial populations such as this, we devised a macro to calculate an age, race and sex adjusted rate. It is now being used to calculate lung and other cancer mortality rates in the state of Ohio.

The variance is estimated assuming that the incident cases follow a Poisson distribution. Therefore, the macro can be used if the incidence rates are less than 10% in the population. The macro also produces age-adjusted rates for each of the race-sex combinations, age and sex adjusted rates for each of the races and age and race adjusted rates for men and women. The associated standard errors and 95% confidence intervals are produced for each of the rates.

The macro was originally written using SAS® version 8.2. The output is produced in both HTML and rich-text format using the SAS output delivery system.

DEFINITIONS:

Incidence rate is the number of new cases of a disease or event in a given period of time divided by the number at risk. The denominator is the number of person years.

Mortality rate is the proportion of the population dying from a particular disease during a given time period. The numerator is the number of persons dying from the disease and the denominator is the appropriate person-years.

This macro calculates the Incidence or Mortality rates to be expressed as per 100000. However, rates per 100, 1000 or any other unit can be calculated by slight modification of the macro.

The Race, Sex and Age specific rate is defined as:

$$R_{ijk} = (C_{ijk} / P_{ijk}) * 100000$$

Where C_{ijk} = Number of incidences in Sex 'i', race 'j' and age-group 'k'.

P_{ijk} = Number of person years for Sex 'i', race 'j' and age-group 'k' in the same specified period of time from which the incidences came.

Adjustments of rates are usually done for standardization and make the rates comparable to other populations and years. The adjusted rates take the form of a weighted average of the age, race and sex specific rates where the sum of the weights is unity.

The Race and Sex specific, age-adjusted Rate can be written as

$$RSR_{ij} = \sum_{k=1}^K (N_{ijk} / N_{...}) * R_{ijk}$$

Where N_{ijk} is the number of persons of sex 'i', race 'j' and age-group 'k' in the reference population to which the rates for each of race-sex groups are to be age adjusted.

The age and sex adjusted rates for a particular race are obtained by multiplying the age and sex specific incidence rate for that race by the age and sex specific proportions in the reference adjusting population and summing over all sex and age groups. Similarly the age and race adjusted rates for each sex is obtained by multiplying the age and race specific incidence rate for that sex by the age and race specific proportions in the reference adjusting population and summing over all race and age groups. Finally, the overall age, race and sex adjusted rates are obtained by multiplying the age, race and sex specific incidence rate by the age, race and sex specific proportions in the reference adjusting population and summing over all age, race and sex groups. Rates can be written notationally as follows:

Race Specific, Age and Sex adjusted Rates:

$$RR_j = \sum_{i=1}^2 \sum_{k=1}^K (N_{ijk} / N_{...}) * R_{ijk}$$

Sex Specific, Age and Race adjusted Rates:

$$RS_i = \sum_{j=1}^2 \sum_{k=1}^K (N_{ijk} / N_{...}) * R_{ijk}$$

Race, Sex and Age adjusted Rates:

$$R = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^K (N_{ijk} / N_{...}) * R_{ijk}$$

Assuming a Poisson distribution for the incident cases, standard errors of the adjusted rates can be estimated as follows. The summation will be over sex (i), race (j) or age (k) as appropriate for the rate under consideration.

If w_{ijk} = weight or the adjusting fraction, then $\sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^K w_{ijk} = 1$. Then the adjusted rate can be

expressed as

$$R = \sum w_{ijk} * R_{ijk} = \sum w_{ijk} * (C_{ijk} / P_{ijk}) * 100000$$

where Σ represents the summation over all three subscripts.

The estimated variance of the adjusted rate is

$$v(Poisson) = \text{var}(R) = \sum \frac{w_{ijk}^2 * C_{ijk} * 100000^2}{P_{ijk}^2} = \sum \frac{w_{ijk}^2 R_{ijk}^2}{P_{ijk}} * 100000$$

The $\sqrt{v(Poisson)}$ is the estimated standard error of the adjusted rate assuming a Poisson distribution.

The 95% confidence interval is then calculated as the adjusted rate plus and minus 1.96 times the standard error. Negative lower limits of the confidence limits were set to zero programmatically.

If the incidences were greater than 10% of the population, one should assume a binomial distribution. The formula for the variance component assuming a binomial distribution can be obtained from the technical report listed in the reference and following the summations as described above, and followed through in the macro.

DATA:

The population for the five county Cincinnati areas by race sex and single year of age was obtained from the CDC/NCHS website. The U.S population estimates for blacks and whites by age and sex in 1990 and 2000 were also obtained from the American Factfinder website. The format of incidence data, denominator population data and the adjustment data are described in the macro itself.

In successive determinations of rates over years, it is necessary to use the same adjusting population, in order to compare the rates. More information on the choice of the adjusting population can be obtained from the publication (2).

OUTPUT:

As an example of output obtained from this macro, the rates of a particular phenotype in the Greater Cincinnati Area are shown in Table 1. The macro also produces a table detailing the number of incidences, incident population and adjusting population. However, that table is not presented here due to space considerations.

CONCLUSION:

In order to calculate overall incidence and mortality rates in populations consisting of two races, it is necessary to adjust across race in addition to age and sex. The formula for the calculation of the variance depends on the assumption of the distribution of the incidences. To keep the calculation of rates in successive years comparable, the same adjusting population must be used for each determination.

As the populations become more diverse, it may become necessary to adjust for more than two races. In future, we hope to expand this macro to calculate rates in multi-racial populations.

REFERENCES:

1. E.J. Bergstralh, K.P. Offord, C.P. Chu, C.M. Beard, W.M. O'Fallon, and L.J. Melton III (1992). Calculating Incidence, Prevalence and Mortality Rates for Olmsted County, Minnesota: An Update. Technical Report Series No: 49, Department of Health Sciences Research, Mayo Clinic Rochester, Minnesota.
2. D.L. Hoyert, R.N. Anderson (2001). Age-adjusted Death Rates: Trend Data Based on the Year 2000 Standard Population. National Vital Statistics Reports. 49:9 .

Table 1: Rates of Phenotype

Type of Rate	Rate 100000	Std Error	95% Lower CI	95% Upper CI
Age Adj:WF	23.09	1.08	20.97	25.22
Age Adj:WM	27.27	1.46	24.40	30.13
Age Adj:BF	39.50	3.78	32.10	46.90
Age Adj:BM	51.68	5.17	41.54	61.82
Age Adj:Female	25.37	1.07	23.28	27.46
Age and Race Adj:Female	24.87	1.05	22.81	26.93
Age Adj:Male	30.36	1.43	27.55	33.18
Age and Race Adj:Male	30.29	1.44	27.47	33.11
Age Adj:White	24.98	0.87	23.28	26.69
Age and Sex Adj:White	25.04	0.87	23.33	26.76
Age Adj:Black	45.26	3.11	39.15	51.36
Age and Sex Adj:Black	45.99	3.17	39.78	52.20
Age Adj:Total	27.72	0.86	26.03	29.41
Age Race and Sex Adj:Total	27.44	0.86	25.76	29.12

ACKNOWLEDGEMENT:

This work was supported in part by grants from NCI (P01-CA096964) and NINDS (R01-NS36695).

CONTACT INFORMATION:

Padmini Sekar
Department of Environmental Health
Kettering Building Room 129
3223 Eden Avenue ML0056
University of Cincinnati Medical Center
Cincinnati, OH-45267-0056
Telephone: (513) 558-1714
Email: sekarpc@ucmail.uc.edu

```
/****RsaRateCal.sas Macro to calculate the race, sex and age adjusted  
macro******/
```

```
/******
```

Call statement:

```
%ratecal(cdata=, agevar=, b_age=,  
          pop=, popyrs=,  
          maxage=, adj=, adj_pop=, macd=, maca=,  
          outdata=, dtype=,rtff=, htmlf=);
```

All input parameters are keyword. They are described below:

CDATA =SAS dataset containing one observation for each incidence case. This dataset must have a character variable named SEX with levels 'M' and 'F', RACE variable with levels 'W' and 'B', and also the age variable to be described next.

AGEVAR=variable defining the integer age at time of diagnosis (onset), prevalence, or death as appropriate for the cases.

B_AGE =String of ordered numbers, a1 a2 a3 etc., used to formulate the desired age groupings for the incidence rates.

Specifying 0 40 60 80 100 implies age intervals of 0-39, 40-59, 60-79, 80-99 and 100 to MAXAGE, where MAXAGE is defined below. Note that, if the smallest age in the string is greater than zero, then any cases and denominator populations less than that age will be excluded from all rate and adjusted rate calculations. Users must exercise caution in interpreting findings when certain ages have been excluded.

POP = Name of the population (denominator) SAS dataset to be used for the incidence rates. Presently the data that is being used is 'Denom4', which has the five county population for years between 1990-2003. The variables pertaining to population for each of the years have a prefix of 'p' i.e., the population for 1993 is in the variable p1993. The Denom4 has age in integers in the variable 'age'. However, the population age 85 and above is presented as 85 years of age. Therefore one may not be able to have any bifurcation between 85 and the maximum age in the B_age macro variable. It also has race and sex which are defined exactly as in CDATA.

POPYRS= variable names(or functions thereof) in POP dataset containing the population counts for each calendar year. Examples are given below:

POPYRS=P1990, rates based on 1990 population.

POPYRS=P1990+P1991+P1992, rates based on the sum of the population from 1990 to 1992, inclusive.

POPYRS=sum(of P1980-P1989), rates based on sum of the population from 1980 to 1989.

POPYRS=((7/12)*p1998)+p1999+p2000+((7/12)*p2001)+((9/12)*p2002)
 Includes seven months each of the years 1998 and 2001,
 nine months of 2002 and the whole of 1999 and 2000. Rates
 based on the sum of those populations. This allows you to use data
 collected in only a part of the year.

MAXAGE=Maximum integer age to use for the incidence analysis.
 The minimum age is the lowest value of the B_AGE string.

ADJ = Name of the data set with adjusting population. The adjusting data
 set should have Race and Sex defined exactly as in CDATA and POP. The age
 could be integer or categorized in small intervals. Presently the
 data set used for this purpose is either 'Adjrs' (for adjustment with Year
 2000 population) or 'Adjrs90' (for adjustment with Year 1990 population)

ADJ_POP=Name of the variable in the 'ADJ' that has the adjusting population. The
 one presently used in Adjrs is US2000. The variable in Adjrs90 is US1990.
 These are in thousands.

OUTDATA=Name of the data set with the calculated incidence rates.

DTYPE =The name or the type of event for which we are calculating the incidence rate.
 The string you specify goes into the title of the output. If you specify
 'ICH',
 the titles become 'Cases of ICH, Incidence Population and Adjusting
 Population'
 and 'Rates of ICH'.

rtff =Name of the .rtf file created with the output. The path should also be
 specified.
 Besides looking nicer, it opens up without alignment problems in Word, even
 if the receiver does not have SAS on their computer.

htmlf =Name of the .html file created with the output. The path should also be
 specified.

Other Conditions:

The libname rdat should be defined and should contain the denominator as well as the
 adjustment data set

The data set referring to cdata should be defined and in the work directory.

 ***/

```
%MACRO AGEGRP(_age);
  %LET BE=X;
  %LET I=1;
  %let error2= ;
  LENGTH AGEGRP $ 6 age_gp $5;
  %DO %UNTIL (&BE= );

    %LET BI=%SCAN(&B_AGE,&I);
    %LET BE=%SCAN(&B_AGE,&I+1);
    %IF &I=1 %THEN %LET MINAGE=&BI;
    %LET BEE=%EVAL(&BE-1);

    %IF &BE^= %THEN %DO;
      %IF &BE<=&BI %THEN %DO;
        %let error2=1;
        %goto error2;
      %END;
      IF &BI <=&_age< &BE THEN AGEGRP="&BI-&BEE";
    %END;

    %IF &BE= %THEN %DO;
      IF &_age >=&BI THEN AGEGRP="&BI-&MAXAGE";
    %END;
  %END;
```

```

        %END;

        %LET I=%EVAL(&I+1);
    %END;
    agegp=left(agegp);
    IF &MINAGE LE &_AGE LE &MAXAGE ;
%error2:
    %if &error2=1 %then %do;
        %PUT ERROR: Age cutpoints not increasing;
        age_gp=.;
    %end;

    %let len= %length(&b_age);
    %let len2=%eval(&len-1);
    %let mcut=%substr(&b_age,&len2,2);
    if agegp="&mcut-&maxage" then agegp="&mcut._ab";

    age_gp=trim(left(agegp));
    drop agegp;
%MEND AGEGRP;

%macro ratecal(cdata=, agevar=, b_age=, pop=, popyrs=, maxage=,
               adj=, adj_pop=, outdata=, dtype=,rtff=, htmlf=);

data &cdata; set &cdata;
    %agegrp(&agevar);
keep race sex &agevar age_gp;
run;

/*
*****
* COMMENT 1: THE FOLLOWING CODE CALCS.
*           SEX*AGE_GP FREQUENCY, RACE*SEX*AGE_GP FREQUENC FOR
*           THE INCIDENCE CASES
*****
*/

PROC FREQ DATA=&cdata;
    TABLES RACE*SEX*AGE_GP / NOPRINT OUT=_OUTRC;
PROC FREQ DATA=&cdata;
    TABLES SEX*AGE_GP / NOPRINT OUT=_OUTC;
PROC FREQ DATA=&cdata;
    TABLES RACE*AGE_GP/ NOPRINT OUT=_OUTR;
PROC FREQ DATA=&cdata;
    TABLES AGE_GP / NOPRINT OUT=_OUTCT;
run;

/*
*****
* COMMENT 2: THE FOLLOWING CODE CALCS.
*           RACE*SEX*AGE_GP FREQUENCY AND SEX*AGE_GP FREQUENCY IN
*           THE DENOMINATOR DATA-For the entire period
*           of the data-both whites and blacks
*****
*/

data denpop; set rdat.&pop;
    %agegrp(&agevar);
    cum_pop=&popyrs;
run;

PROC FREQ DATA=denpop;
    TABLES RACE*SEX*AGE_GP / NOPRINT OUT=_OUTRI;
    WEIGHT CUM_POP;
PROC FREQ DATA=denpop;
    TABLES SEX*AGE_GP / NOPRINT OUT=_OUTI;
    WEIGHT CUM_POP;
PROC FREQ DATA=denpop;
    TABLES RACE*AGE_GP / NOPRINT OUT=_OUTCI;

```

```

        WEIGHT CUM_POP;
PROC FREQ DATA=denpop;
    TABLES          AGE_GP /  NOPRINT OUT=_OUTIT;
    WEIGHT CUM_POP;
run;

DATA _CASES  (RENAME=(COUNT=C_COUNT));  SET _OUTCT _OUTC _OUTR _OUTRC;

DATA _INC_POP(RENAME=(COUNT=I_COUNT));  SET _OUTIT _OUTCI _OUTI _OUTRI;
RUN;

/*
*****
* COMMENT 3 : Code for adjusted incidence rates follows
*
*
*****
*/

Data _ADJ; set rdat.&adj;
    %agegrp(&agevar);
keep race age_gp sex &adj_pop;
run;

*** sum adjusted populations by age groupings, gender and race;

PROC FREQ DATA=_ADJ;
    TABLES RACE*SEX*AGE_GP /  NOPRINT OUT=_OUTRA;
    WEIGHT &adj_pop;

PROC FREQ DATA=_ADJ;
    TABLES      SEX*AGE_GP /  NOPRINT OUT=_OUTA;
    WEIGHT &adj_pop;

PROC FREQ DATA=_ADJ;
    TABLES      RACE*AGE_GP /  NOPRINT OUT=_OUTCA;
    WEIGHT &adj_pop;

PROC FREQ DATA=_ADJ;
    TABLES      AGE_GP /  NOPRINT OUT=_OUTAT;
    WEIGHT &adj_pop;

DATA _ADJ_POP(RENAME=(COUNT=A_COUNT));  SET _OUTAT _OUTCA _OUTA _OUTRA;
RUN;

proc sort data=_inc_pop; by race sex age_gp; run;
proc sort data=_adj_pop; by race sex age_gp; run;
proc sort data=_cases; by race sex age_gp; run;

DATA _T;
    MERGE _INC_POP _ADJ_POP _CASES; BY RACE SEX AGE_GP;
    IF C_COUNT=. THEN C_COUNT=0;
    IF I_COUNT=. THEN I_COUNT=0;
    IF A_COUNT=. THEN A_COUNT=0;
    CASES=C_COUNT;
    POPN=I_COUNT;
    ADJ_POP=A_COUNT;
    KEEP RACE SEX AGE_GP CASES POPN ADJ_POP;
run;

DATA _WF; SET _T;          IF SEX='F' and RACE='W';
DATA _WM; SET _T;          IF SEX='M' and RACE='W';
DATA _BF; SET _T;          IF SEX='F' and RACE='B';
DATA _BM; SET _T;          IF SEX='M' and RACE='B';
DATA _F; SET _T;           IF SEX='F' and RACE=' ' ; RACE='X';
DATA _M; SET _T;           IF SEX='M' and RACE=' ' ; RACE='X';
DATA _W; SET _T;           IF SEX=' ' AND RACE='W'; SEX='T';
DATA _BL; SET _T;          IF SEX=' ' AND RACE='B'; SEX='T';
DATA _B; SET _T;           IF SEX=' ' AND RACE=' ' ; SEX='T'; RACE='X';

RUN;

```

```

DATA _ALL; MERGE      _WF(RENAME=(CASES=WF_C POPN=WF_P ADJ_POP=WF_AP))
                     _WM(RENAME=(CASES=WM_C POPN=WM_P ADJ_POP=WM_AP))
                     _BF(RENAME=(CASES=BF_C POPN=BF_P ADJ_POP=BF_AP))
                     _BM(RENAME=(CASES=BM_C POPN=BM_P ADJ_POP=BM_AP))
                     _F(RENAME=(CASES=F_C POPN=F_P ADJ_POP=F_AP))
                     _M(RENAME=(CASES=M_C POPN=M_P ADJ_POP=M_AP))
                     _W(RENAME=(CASES=W_C POPN=W_P ADJ_POP=W_AP))
                     _BL(RENAME=(CASES=B_C POPN=B_P ADJ_POP=B_AP))
                     _B(RENAME=(CASES=T_C POPN=T_P ADJ_POP=T_AP))
                     ;
                     BY AGE_GP;
WF_I=(WF_C/WF_P)*100000;
WM_I=(WM_C/WM_P)*100000;
BF_I=(BF_C/BF_P)*100000;
BM_I=(BM_C/BM_P)*100000;
F_I=(F_C/F_P)*100000;
M_I=(M_C/M_P)*100000;
W_I=(W_C/W_P)*100000;
B_I=(B_C/B_P)*100000;
T_I=(T_C/T_P)*100000;

*Adjusted rate components;
aarg_wf=wf_i*t_ap;
aarg_wm=wm_i*t_ap;
aarg_bf=bf_i*t_ap;
aarg_bm=bm_i*t_ap;

aarsp_wf=wf_i*f_ap;
aarsp_wm=wm_i*m_ap;
aarsp_bf=bf_i*f_ap;
aarsp_bm=bm_i*m_ap;

aarsr_wf=wf_i*w_ap;
aarsr_wm=wm_i*w_ap;
aarsr_bf=bf_i*b_ap;
aarsr_bm=bm_i*b_ap;

aarrs_wf=wf_i*wf_ap;
aarrs_wm=wm_i*wm_ap;
aarrs_bf=bf_i*bf_ap;
aarrs_bm=bm_i*bm_ap;

agar_f=f_i*t_ap;
agar_m=m_i*t_ap;

agar_w=w_i*t_ap;
agar_b=b_i*t_ap;

arrs_t=t_i*t_ap;
asar_t=.;

*ADJUSTED RATE VARIANCE COMPONENTS;
varg_wf=(t_ap**2)*(wf_i/wf_p)*100000;
varg_wm=(t_ap**2)*(wm_i/wm_p)*100000;
varg_bf=(t_ap**2)*(bf_i/bf_p)*100000;
varg_bm=(t_ap**2)*(bm_i/bm_p)*100000;

varsp_wf=(f_ap**2)*(wf_i/wf_p)*100000;
varsp_wm=(m_ap**2)*(wm_i/wm_p)*100000;
varsp_bf=(f_ap**2)*(bf_i/bf_p)*100000;
varsp_bm=(m_ap**2)*(bm_i/bm_p)*100000;

varsr_wf=(w_ap**2)*(wf_i/wf_p)*100000;
varsr_wm=(w_ap**2)*(wm_i/wm_p)*100000;
varsr_bf=(b_ap**2)*(bf_i/bf_p)*100000;
varsr_bm=(b_ap**2)*(bm_i/bm_p)*100000;

varrs_wf=(wf_ap**2)*(wf_i/wf_p)*100000;
varrs_wm=(wm_ap**2)*(wm_i/wm_p)*100000;
varrs_bf=(bf_ap**2)*(bf_i/bf_p)*100000;

```

```

varrs_bm=(bm_ap**2)*(bm_i/bm_p)*100000;

vgar_f=(t_ap**2)*(f_i/f_p)*100000;
vgar_m=(t_ap**2)*(m_i/m_p)*100000;

vgar_w=(t_ap**2)*(w_i/w_p)*100000;
vgar_b=(t_ap**2)*(b_i/b_p)*100000;

vrrs_t=(t_ap**2)*(t_i/t_p)*100000;

format aarg_wf--arrrs_t 10.4;

keep age_gp wf_c wm_c bf_c bm_c f_c m_c w_c b_c t_c wf_p wm_p bf_p bm_p f_p m_p
      w_p b_p t_p wf_i wm_i bf_i bm_i f_i m_i w_i b_i t_i
      wf_ap wm_ap bf_ap bm_ap f_ap m_ap w_ap b_ap t_ap aarg_wf--vrrs_t;

RUN;

proc means data=_all noprint;
  var wf_c wm_c bf_c bm_c f_c m_c w_c b_c t_c wf_p wm_p bf_p bm_p f_p m_p w_p b_p t_p
      wf_ap wm_ap bf_ap bm_ap f_ap m_ap w_ap b_ap t_ap aarg_wf--vrrs_t;
output out=_adj sum=wf_c wm_c bf_c bm_c f_c m_c w_c b_c t_c wf_p wm_p bf_p bm_p f_p
m_p
      w_p b_p t_p wf_ap wm_ap bf_ap bm_ap f_ap m_ap w_ap b_ap t_ap
aarsp_bm
      aarsr_wf aarsr_wm aarsr_bf aarsr_bm aarrs_wf aarrs_wm aarrs_bf
      aarrs_bm agar_f agar_m agar_w agar_b arrs_t asar_t
varsp_bm
      varsr_wf varsr_wm varsr_bf varsr_bm varrs_wf varrs_wm varrs_bf
      varrs_bm vgar_f vgar_m vgar_w vgar_b vrrs_t;

run;

DATA _ADJ2; SET _ADJ;
aarg_wf=aarg_wf/t_ap;
serg_wf=sqrt(varg_wf/t_ap**2);
llrg_wf=aarg_wf-1.96*serg_wf; if .z<llrg_wf<0 then llrg_wf=0;
ulrg_wf=aarg_wf+1.96*serg_wf;

aarg_wm=aarg_wm/t_ap;
serg_wm=sqrt(varg_wm/t_ap**2);
llrg_wm=aarg_wm-1.96*serg_wm; if .z<llrg_wm<0 then llrg_wm=0;
ulrg_wm=aarg_wm+1.96*serg_wm;

aarg_bf=aarg_bf/t_ap;
serg_bf=sqrt(varg_bf/t_ap**2);
llrg_bf=aarg_bf-1.96*serg_bf; if .z<llrg_bf<0 then llrg_bf=0;
ulrg_bf=aarg_bf+1.96*serg_bf;

aarg_bm=aarg_bm/t_ap;
serg_bm=sqrt(varg_bm/t_ap**2);
llrg_bm=aarg_bm-1.96*serg_bm; if .z<llrg_bm<0 then llrg_bm=0;
ulrg_bm=aarg_bm+1.96*serg_bm;

aarrs_f=(aarsr_wf+aarsr_bf)/t_ap;
sers_f=sqrt((varsr_wf+varsr_bf)/t_ap**2);
llrs_f=aarrs_f-1.96*sers_f; if .z<llrs_f<0 then llrs_f=0;
ulrs_f=aarrs_f+1.96*sers_f;

aarrs_m=(aarsr_wm+aarsr_bm)/t_ap;
sers_m=sqrt((varsr_wm+varsr_bm)/t_ap**2);
llrs_m=aarrs_m-1.96*sers_m; if .z<llrs_m<0 then llrs_m=0;
ulrs_m=aarrs_m+1.96*sers_m;

aarrs_w=(aarsp_wf+aarsp_wm)/t_ap;
sers_w=sqrt((varsp_wf+varsp_wm)/t_ap**2);
llrs_w=aarrs_w-1.96*sers_w; if .z<llrs_w<0 then llrs_w=0;

```

```

ulrs_w=aarrs_w+1.96*sers_w;

aarrs_b=(aarsp_bf+aarsp_bm)/t_ap;
sers_b=sqrt((varsp_bf+varsp_bm)/t_ap**2);
llrs_b=aarrs_b-1.96*sers_b; if .z<llrs_b<0 then llrs_b=0;
ulrs_b=aarrs_b+1.96*sers_b;

agar_f=agar_f/t_ap;
sear_f=sqrt(vgar_f/t_ap**2);
llrg_f=agar_f-1.96*sear_f; if .z<llrg_f<0 then llrg_f=0;
ulrg_f=agar_f+1.96*sear_f;

agar_m=agar_m/t_ap;
sear_m=sqrt(vgar_m/t_ap**2);
llrg_m=agar_m-1.96*sear_m; if .z<llrg_m<0 then llrg_m=0;
ulrg_m=agar_m+1.96*sear_m;

agar_w=agar_w/t_ap;
sear_w=sqrt(vgar_w/t_ap**2);
llrg_w=agar_w-1.96*sear_w; if .z<llrg_w<0 then llrg_w=0;
ulrg_w=agar_w+1.96*sear_w;

agar_b=agar_b/t_ap;
sear_b=sqrt(vgar_b/t_ap**2);
llrg_b=agar_b-1.96*sear_b; if .z<llrg_b<0 then llrg_b=0;
ulrg_b=agar_b+1.96*sear_b;

arrs_t=arrs_t/t_ap;
serrs_t=sqrt(vrrs_t/t_ap**2);
lrrs_t=arrs_t-1.96*serrs_t; if .z<lrrs_t<0 then lrrs_t=0;
urrs_t=arrs_t+1.96*serrs_t;

asar_t=(aarrs_wf+aarrs_wm+aarrs_bf+aarrs_bm)/t_ap;
sear_t=sqrt((varrs_wf+varrs_wm+varrs_bf+varrs_bm)/t_ap**2);
lsar_t=asar_t-1.96*sear_t; if .z<lsar_t<0 then lsar_t=0;
usar_t=asar_t+1.96*sear_t;

drop aarsp_wf aarsp_wm aarsp_bf aarsp_bm aarsr_wf aarsr_wm aarsr_bf aarsr_bm
      varsp_wf varsp_wm varsp_bf varsp_bm varsr_wf varsr_wm varsr_bf varsr_bm;
run;

data &outdata; set _all(drop=aarsp_wf aarsp_wm aarsp_bf aarsp_bm aarsr_wf aarsr_wm
aarsr_bf
      aarsr_bm varsp_wf varsp_wm varsp_bf varsp_bm varsr_wf varsr_wm varsr_bf
varsr_bm)
      _adj2(in=inj);
if inj then age_gp='TOTAL';
if inj then do;
wf_i=(wf_c/wf_p)*100000;
wm_i=(wm_c/wm_p)*100000;
bf_i=(bf_c/bf_p)*100000;
bm_i=(bm_c/bm_p)*100000;
w_i=(w_c/w_p)*100000;
b_i=(b_c/b_p)*100000;
m_i=(m_c/m_p)*100000;
f_i=(f_c/f_p)*100000;
t_i=(t_c/t_p)*100000;
end;
if age_gp ne 'TOTAL' then do;
aarg_wf=.; aarg_wm=.; aarg_bf=.; aarg_bm=.; aarrs_f=.; aarrs_m=.;
aarrs_w=.; aarrs_b=.; agar_f=.; agar_m=.; agar_w=.; agar_b=.;
arrs_t=.; asar_t=.;
end;
keep age_gp wf_c wm_c bf_c bm_c w_c b_c f_c m_c t_c wf_p wm_p bf_p bm_p
      w_p b_p f_p m_p t_p wf_i wm_i bf_i bm_i w_i b_i f_i m_i t_i
      wf_ap wm_ap bf_ap bm_ap w_ap b_ap f_ap m_ap t_ap
      aarg_wf aarg_wm aarg_bf aarg_bm aarrs_f aarrs_m
      aarrs_w aarrs_b agar_f agar_m agar_w agar_b
      aarrs_w aarrs_b aarrs_f aarrs_m arrs_t asar_t serg_wf--usar_t;
format wf_i wm_i bf_i bm_i w_i b_i f_i m_i t_i 9.3;

```

```

run;

data wf; set &outdata;
  keep age_gp wf_c wf_p wf_ap wf_i;
run;

data wf; set wf;
  rename wf_c=Numcase
         wf_p=Popul
         wf_ap=AdjPop
         wf_i=Incidnce;
  racesex='WF';
run;

data wm; set &outdata;
  keep age_gp wm_c wm_p wm_ap wm_i;
run;

data wm; set wm;
  rename wm_c=Numcase
         wm_p=Popul
         wm_ap=AdjPop
         wm_i=Incidnce;
  racesex='WM';
run;

data bf; set &outdata;
  keep age_gp bf_c bf_p bf_ap bf_i;
run;

data bf; set bf;
  rename bf_c=Numcase
         bf_p=Popul
         bf_ap=AdjPop
         bf_i=Incidnce;
  racesex='BF';
run;

data bm; set &outdata;
  keep age_gp bm_c bm_p bm_ap bm_i;
run;

data bm; set bm;
  rename bm_c=Numcase
         bm_p=Popul
         bm_ap=AdjPop
         bm_i=Incidnce;
  racesex='BM';
run;

data pop2; set wf wm bf bm;
rs=racesex;
rcsx=racesex;
label
Numcase='CASES'
Popul='POPN'
AdjPop='ADJ_POP'
Incidnce='INCIDENCE PER 100000'
;
run;

%macro splita(da,tp,ra,sev,lr,ur);
  data &da; set &outdata;
  if age_gp='TOTAL';
  attrib typ length=$30 label='Type of Rate';
  typ="&tp";
  keep typ &ra &sev &lr &ur;
run;

data &da; set &da;
  rename &ra=IncRate

```

```

        &sev=StdEr
        &lr =LowerCI
        &ur =UpperCI;
run;
%mend;

%splita(a, Age Adj:WF, aarg_wf, serg_wf, llrg_wf, ulrg_wf);
%splita(b, Age Adj:WM, aarg_wm, serg_wm, llrg_wm, ulrg_wm);
%splita(c, Age Adj:BF, aarg_bf, serg_bf, llrg_bf, ulrg_bf);
%splita(d, Age Adj:BM, aarg_bm, serg_bm, llrg_bm, ulrg_bm);

%splita(e, Age Adj:Female, agar_f, sear_f, llrg_f, ulrg_f);
%splita(f, Age and Race Adj:Female, aarrs_f, sers_f, llrs_f, ulrs_f);
%splita(g, Age Adj:Male, agar_m, sear_m, llrg_m, ulrg_m);
%splita(h, Age and Race Adj:Male, aarrs_m, sers_m, llrs_m, ulrs_m);

%splita(i, Age Adj:White, agar_w, sear_w, llrg_w, ulrg_w);
%splita(j, Age and Sex Adj:White, aarrs_w, sers_w, llrs_w, ulrs_w);
%splita(k, Age Adj:Black, agar_b, sear_b, llrg_b, ulrg_b);
%splita(l, Age and Sex Adj:Black, aarrs_b, sers_b, llrs_b, ulrs_b);

%splita(m, Age Adj:Total, arrs_t, serrs_t, llrs_t, urrs_t);
%splita(n, Age Race and Sex Adj:Total, asar_t, sear_t, lsar_t, usar_t);

data rprnt; set a b c d e f g h i j k l m n;
label Incrate='Rate Per 100000'
      StdEr='Std Error'
      LowerCI='95% Lower Conf Limit'
      UpperCI='95% Upper Conf Limit' ;
run;

option orientation=landscape;

ods rtf file="&rtff";
ods html file="&htmlf";

proc report nowindows data=pop2;
  column age_gp racesex, Numcase rcsx, popul rs, adjpop;
  define age_gp /group format=$5. width=5 'AgeGr';
  define racesex/across width=9 '_CASES_';
  define numcase/format=4.0 width=5;
  define rcsx/across width=9 '_Population_';
  define popul/ format=9.0 width=10;
  define rs/across width=9 '_Adj Pop in 1000s_';
  define adjpop/ format=7.0 width=8;
  rbreak before/ol;
  rbreak after/ol;
  title "Cases of &dtype, Incident Population and Adjusting Population";
run;

proc report nowindows data=rprnt;
  column typ Incrate Stder lowerci upperci;
  format Incrate Stder lowerci upperci 9.2;
  rbreak before/ol;
  rbreak after/ol;
  title "Rates of &dtype";
run;
ods html close;
ods rtf close;

title;

%mend ratecal;

```