

```

libname d "C:\Documents and Settings\David\My documents\MDM\DCCTdata5";
*Data on proportions of diets from protein, carbohydrates, fiber,PUFA, MUFA, SFA, total fat, alcohol
  from >168 countries
  from FAO;
data BMIFAOMN;
  set d.BMIFAOMN;

*Data on cardiovascular disease risks from >168 countries from WHO, Institute of Health Metrics
  and Evaluation and International Diabetes Federation;
data BMIFAOIm;
  set d.BMIFAOIm;
*Combining the data;
data d.source2;
merge BMIFAOIm BMIFAOMN;
by mask_pat;
  drop _type_ _freq_;

label
calave = mean calories
proave = mean protein g
tcarbave =mean total carbs g
dfave = mean diet fiber g
tfatave = mean total fat g
sfatave = mean sat fat g
mfatave=mean mono fat g
pfatave =mean poly fat g
MK_4=vitamin K2 menaquinone-4
K1=vitamin K1 phylloquinone;

*Macronutrients by percentage of caloric intake;
if calave > 0 then pprot=(proave*4 /calave) * 100;
if calave > 0 then pcarbs=(tcarbave*4 /calave) * 100;
if calave > 0 then dfp1000=dfave / (calave/ 1000);
if calave > 0 then ptfat=(tfatave*9/calave) * 100;
if calave > 0 then psfa=(sfatave*9/calave) *100;
if calave > 0 then pmufa=(mfatave*9/calave) * 100;
if calave > 0 then ppufa=(pfatave*9/calave) * 100;
if calave > 0 then ptransft=(transfat*9/calave) *100;
if (calave > 0 and dfave > 0) then carb_DF=tcarbave/dfave;
if (calave > 0 and pfatave>0) then TF_PUFA=tfatave/pfatave;
if calave > 0 then MK_41000=MK_4/(calave/1000);
if calave > 0 then MK_42000=2*MK_4/(calave/1000);
if calave > 0 then MK513_1000=MK_5_13/(calave/1000);
if calave > 0 then MK_n=MK_4 +MK_5_13;
if calave > 0 then MKn_1000=MK_41000 + MK513_1000;
if calave > 0 then MKn_2000=(MK_41000 + MK513_1000)*2;
if calave > 0 then K11000=K1/(calave/1000);
if calave > 0 then K12000=2*K1/(calave/1000);

if calave > 0 then animala2=animalav/(calave/1000);
if calave > 0 then vegeta2=vegetav/(calave/1000);
if calave > 0 then dairyav=butterav+creamav+milkav+cheeseav;
if calave > 0 then dairya2=dairyav/(calave/1000);

if calave > 0 then Buttera2=Butterav/(calave/1000);
if calave > 0 then Cheesea2=Cheeseav/(calave/1000);

```

```
if calave > 0 then Creama2=Creamav/(calave/1000);
if calave > 0 then Milka2=Milkav/(calave/1000);
if calave > 0 then eggsa2=eggsav/(calave/1000);
if calave > 0 then meata2=meatav/(calave/1000);
if calave > 0 then cowa2=cowav/(calave/1000);
if calave > 0 then muttona2=muttonav/(calave/1000);
if calave > 0 then piga2=pigav/(calave/1000);
if calave > 0 then pouлта2=poultav/(calave/1000);
if calave > 0 then fisha2=fishav/(calave/1000);
if calave > 0 then offala2=offalav/(calave/1000);
if calave > 0 then cereala2=cerealav/(calave/1000);
if calave > 0 then vegoila2=vegoilav/(calave/1000);
if calave > 0 then starcha2=starchav/(calave/1000);
if calave > 0 then sugara2=sugarav/(calave/1000);
if calave > 0 then pulsesa2=pulsesav/(calave/1000);
if calave > 0 then treenta2=treentav/(calave/1000);
if calave > 0 then oila2=oilav/(calave/1000);
if calave > 0 then vegoila2=vegoilav/(calave/1000);
if calave > 0 then vegesa2=vegesav/(calave/1000);
if calave > 0 then fruitsa2=fruitsav/(calave/1000);
```

label

pprot=% protein

ptcarbs=% total carbs

dfp1000=dietary fiber g per 1000 kc

ptfat=% total fat

psfa=% SFA

pmufa=% MUFA

ppufa= % PUFA

ptransft=% transfat

carb_df=carb-fiber ratio

TF_PUFA=Total Fat-PUFA ratio

animala2=animal per 1000kcs

vegeta2=plant-based food per 1000kcs

Buttera2=Butter per 1000kcs

Cheesea2=Cheese per 1000kcs

Creama2=Cream per 1000kcs

Milka2=Milk per 1000kcs

eggsa2=eggs per 1000kcs

cowa2=beef per 1000kcs

muttona2=mutton per 1000kcs

piga2=pork per 1000kcs

pouлта2=poultry per 1000kcs

meata2=meat per 1000kcs

fisha2=fish per 1000kcs

offala2=offal per 1000kcs

cereala2=cereal per 1000kcs

vegoila2=vegoil per 1000kcs

starcha2=starch per 1000kcs

sugara2=sugar per 1000kcs

pulsesa2=pulses per 1000kcs

treenta2=treent per 1000kcs

oila2=oil per 1000kcs

vegesa2=veges per 1000kcs

fruitsa2= fruit per 1000kcs

```

K11000=K1/1000kcs
MK_41000=MK-4/1000 kcs
MK_42000=2*MK_4 per 2000kcs
MK513_1000=MK_5 to MK_13
MK_n=all menaquinones MK5-13
MKn_1000=all menaquinones MK5-13 per 1000kcs
MKn_2000=all menaquinones MK5-13 per 2000kcs
K11000=vitamin K1 per 1000kcs
K12000=vitamin K1 per 2000kcs
;
*Imputing lack of exercise with WHO data (NoPE2010) using a multiple regression formula
(NOPE2010=biometrics, food commodities, and gender). "3.7" is the least insufficient
physical activity in the available data so 3.7 was set as the floor for insufficient
physical activity;

NoPe10 = NoPe2010;
NoPef= (sugara2*0.04138
        +DMPprev15*0.31316
        +FEMALE*7.93392
        -AirPoll*0.10401)* 1.7 +8.5
;

if NoPe2010 = . then Nope10 = NoPef;
if noPe10 < 3.7 then Nope10= 3.7;

label
NoPe2010=insuff PA WHO 2010
NoPef=formula modeling insuff PA
NoPe10=insuff PA WHO 2010 + imputed
;

*Imputing tobacco use (smoke:fraction 0-1 of tobacco users) with multiple regression
with the variables: calories, alcohol g, SBP, and gender for 2013 data;

smokef=-FEMALE*0.18499
        + alc2010*0.00321
;
smokef2= ((-FEMALE*0.18499
        + alc2010*0.00321)*0.67029
        +calave*0.00006163)*1.68 + .044
;
smoke = smoke13;
if smoke13 = . then smoke = smokef2;
if smoke < 0 then smoke = 0;

label
smoke =smoking (yes is 1,no is 0)+imputed
smokef2 = formula to impute % smoking adding gender
;

*formula to impute values for childhood mortality;
mort5f= (-bmi2014 *6.45229 + MBP2014 *1.58036 -animalav *0.06428) * 1.7 +25;

if mort5f < 3 then mort5f = 3;
mort5yrx = mort5yrs;

```

```

if mort5yrx = . then mort5yrx = mort5f;

label
mort5yrs=childhood mortality WHO
mort5f=formula to impute childhood mortality
mort5yrx=childhood mortality + imputed
;
*formula to impute values for years of education;
edu13 = edu2013;
edu2013f = (bmi2014*0.14166 + GDP2009 *0.00004155 + calave * 0.00097468 - mort5yrx*0.04225)* 1.279 + 1
;
if edu2013 = . then edu13 = edu2013f;
if (edu2013 = . and edu13 ne .) then edu13x = edu2013f;
;
label
edu2013f=formula to impute values for years of education
edu13=years education + imputed values
edu13x=formula values for year of education when NA
;

*Cardiovascular disease risk calculation: (CVD15_64t) food % of kcals, MCG/1000kcal(K1 and K2),
and alcohol (g consumed WHO rather than g available FAO), (CVD15_64u) converting to g, MCG,
(CVD15_64v) adding MK-4 per 1000 kcs, (CVD15_64w) changing MK-4 per 1000 kcs to MK-4, (CVD15_64x)adding
factors except air pollution, (CVD15_64y)adding air pollution.(n=336, R2=0.4906);

if calave > 0 then CVD15_64t=alc2010*56.21597
- MKn_1000*114.47155
;

if calave > 0 then CVD15_64u=alc2010*56.21597
- MK_n*114471.55/calave
;

if cvd15_64 ne . then CVD15_64v= (alc2010*56.21597
- MK_n*114471.55/calave)* 0.22354
+ smoke*2255.80139
+ MBP2014*36.85500
- GDP2009*0.01437
+ AirPoll*8.33110
;

if cvd15_64v < 155.5 then cvd15_64v = 155.5;

*Wealthiest cohorts formula for cardiovascular disease risk calculation related code--CVD15_64x:
CVD15_64x: BMI, early childhood mortality, and tobacco (weaker but significant variables),
CVD15_64y: adding other risk factors.(n=70, R2=0.70.16);

CVD15_64x= bmi2014*100.18480
+ mort5yrx*67.78103
+ smoke *1534.93976
;

CVD15_64y=(bmi2014*100.18480
+ mort5yrx*67.78103
+ smoke *1534.93976)* 0.52335
- CHOL*442.02313

```

```
+ MBP2014*34.36743
+ DMPPrev15*9.42543
```

```
;
```

```
*Formula for BMI with the following stages to the multiple regressions:
```

```
(1) nutrients except MK-4-13_1000, (2) converting to transfat g, and k1 (3) adding MK-41000  
(4) converting MK-4-13_1000 to MK-4-13 (5) add other risk factors (n=336, R2=0.6005)
```

```
;
```

```
if calave > 0 then BMIIf1=sugara2* 0.02265  
+ animala2*0.00614
```

```
;
```

```
if calave > 0 then BMIIf2=(sugarav* 22.65  
+ animalav*6.14)/calave
```

```
;
```

```
if calave > 0 then BMIIf3=((sugarav* 22.65  
+ animalav*6.14)/calave)* 0.30383  
+ NoPe10 *0.02118  
+ DMPPrev15*0.08390  
+ FEMALE*0.61815  
+ edu13*0.13124  
- AirPoll*0.00742
```

```
;
```

```
*Formula for prevalence of diabetes in 40-49 year old from IDF data (n=336 cohorts, R2=0.5596)
```

```
(1) nutrients-sugar and animal products (kcal/1000 kcal), (2) add starches and poultry, which have  
strong univariate correlations (3) convert to kcal, (4) add insufficient physical activity, smoking, a  
childhood mortality, and (5) add BMI the dominant variable;
```

```
DMPf1= animala2*0.01714  
+ sugara2*0.09463  
;
```

```
if calave > 0 then DMPf2= (animala2*0.01714  
+ sugara2*0.09463)*0.43567  
- starcha2*0.01558  
+ pouлта2*0.21895  
;
```

```
if calave > 0 then DMPf3= ((animalav*17.14  
+ sugarav*94.63)/calave) * 0.43567  
- (starchav*15.58  
- pouлтаv*218.95)/calave  
;
```

```
if calave > 0 then DMPf4= (((animalav*17.14  
+ sugarav*94.63)/calave) * 0.43567  
- (starchav*15.58  
- pouлтаv*218.95)/calave)*0.41847  
+ smoke*10.66296  
+ NoPe10*0.21511  
- mort5yrx*0.07625
```

```
;
```

```
if calave > 0 then DMPf5=(((animalav*17.14
+ sugarav*94.63)/calave) * 0.43567
- (starchav*15.58
- poultav*218.95)/calave)*0.41847
+ smoke*10.66296
+ NoPe10*0.21511
- mort5yrx*0.07625)*0.74159
+ bmi2014*1.19515
```

```
;
```

*Formula for systolic BP, derived in stages as above (n=336 cohorts, R2=0.4335);

```
if calave > 0 then MBPF1=alc2010*0.46343
-MKn_1000*0.53096
```

```
;
```

```
if calave > 0 then MBPF2=alc2010*0.46343
-MK_n*530.96/calave
```

```
;
```

```
if calave > 0 then MBPF3= (alc2010*0.46343
-MK_n*530.96/calave) * 0.66476
-GDP2009*0.00004485
+mort5yrx*0.02318
+smoke*5.79960
```

```
;
```

```
if calave > 0 then MBPF4= ((alc2010*0.46343
-MK_n*530.96/calave) * 0.66476
-GDP2009*0.00004485
+mort5yrx*0.02318
+smoke*5.79960) *0.78022
-FEMALE*2.60763
```

```
;
```

*Formula for cholesterol staged with (1) and (2)diet, (3) insufficient physical exercise, (4) other risks (n=336 cohorts, R2=0.79.69);

```
cholf1=animala2*0.00298
+sugara2*0.00145
```

```
;
```

```
if calave > 0 then cholf2=(animalav*2.98
+sugarav*1.45)/calave
```

```
;
```

```
if calave > 0 then cholf3=((animalav*2.98
+sugarav*1.45)/calave)*0.94787
+NoPe10 *0.00428
```

```
;
```

```
if calave > 0 then cholf4=(((animalav*2.98
+sugarav*1.45)/calave)*0.94787
+NoPe10 *0.00428)*0.32000
```

+GDP2009*0.00000810
- mort5yrx*0.00435

;

*Calculation for breakdown of attributed risk for diabetes prevalence (IDF);

label
cvd15_64t=CVD formula nutrient %
CVD15_64u=CVD formula nutrients g-MGM
CVD15_64v=CVD formula add smoke SBP gdp & air pollution

CVD15_64x=CVD formula nutrient %

BMIf1=BMI formula nutrients%
BMIf2=BMI formula nutrients g mcg
BMIf3=BMI formula add others

DMPf1=DM formula macronutrients%
DMPf2=DM formula macronutrients g mcg
DMPf3=DM formula add tobacco + air pollution
DMPf4=DM formula add insufficient PE
DMPf5=DM formula add BMI and child death

MBPf1=SBP formula nutrients%
MBPf2=SBP formula nutrients g mcg
MBPf3=SBP formula add all but gender

cholf1=serum cholesterol formula nutrients%
cholf2=serum cholesterol formula nutrients g mcg
cholf3=serum cholesterol formula add other risks
;
run;
quit;