

Multi-Ethnic Study of Atherosclerosis

The MESA Limited Access Dataset release contains two primary datasets, the baseline examination and the results of the diet questionnaire. The baseline and diet files are separated for easier usage; the unique key MESAID can be used to combine the datasets for analyses that use data from both sources. The baseline file contains the results from many different procedures: anthropometry, MRI, and blood groups among others. The reason variables end with "1" is to indicate the data is from the first (baseline) examination; LAD releases of later examinations will have the appropriate number on the end of the variables. Those variables with names ending in "1c" indicate that they are a computed variable. The definitions to these variables can be found in the file "Created_Variables_Dictionary.doc".

Exclusion from LADs

A few variables were excluded from the dataset if they were deemed to have information that could potentially identify participants (specific dates, location identifiers, and a small number of extraordinarily sparse dichotomous variables such as amputation). The vast majority of data has been made available.

Variables with potentially identifying outliers (height, weight, smoking exposure, etc) were investigated during the preparation of the limited access datasets. When it was determined that outliers could provide identifying information, a minor recoding was performed. Continuous variables had their top 5 values recoded to the 6th highest, as well as their lowest 5 values recoded to the 6th lowest. This was performed in each strata of race (race1c) and gender (gender1).

Sparse data situations for categorical variables could also be potentially identifying in some circumstances (e.g. with an extraordinarily sparse category). For categorical variables, the identifying category was combined with one of the other categories (typically "other").

The number of recodes was kept to a minimum: a full list of the recoded variables follows.

MESA EXAM 1 RECODED VARIABLES



ANTHROPOMETRY	
Variable	Comment
htcm1	race/gender recodes
wtlb1	race/gender recodes
bmi1c	based upon height and weight
bmicat1c	based upon bmi
bsa1c	based upon height and weight

HEALTH AND LIFE	
Variable	Comment
uf2fire1	Recode "sexual orientation", "religion" responses to "Other"
uf2hire1	Recode "sexual orientation", "religion" responses to "Other"
uf2stop1	Recode "sexual orientation", "religion" responses to "Other"
uf2educ1	Recode "sexual orientation", "religion" responses to "Other"
uf2move1	Recode "sexual orientation", "religion" responses to "Other"
uf2nghb1	Recode "sexual orientation", "religion" responses to "Other"

MEDICAL HISTORY	
Variable	Comment
lungcn1	n=9 cases, roll into "Other Cancer" variable
bloodcn1	n=4 cases, roll into "Other Cancer" variable
othcn1	Accept cases from lungcn1 and bloodcn1
heptpd1	n=0 cases, roll into "Unknown Type" variable
heptpe1	n=2 cases, roll into "Unknown Type" variable
heptpu1	Accept cases from heptpd1 and heptpe1
hwt3ylb1	race/gender recodes
wt20lb1	race/gender recodes
wt40lb1	race/gender recodes
pregn1	race/gender recodes, values greater than 12 top-coded to 12
birthn1	race/gender recodes, values greater than 10 top-coded to 10
agebrth1	race/gender recodes
hystage1	race/gender recodes
ovaage1	race/gender recodes
menoage1	race/gender recodes
hrmage1	"99" values to set to missing
hrmqage1	race/gender recodes
abnum1	race/gender recodes
ab5ynum1	14 "99" values to set to missing, race/gender.
tthnum1	race/gender recodes

PERSONAL HISTORY

Variable	Comment
----------	---------

stbth1	US Census Regions, correlate typos against parents' values when possible
ctrybth1	Collapse countries with low occurrence into regions
mstbth1	US Regions
fstbth1	US Regions
nhhldc1	race/gender recodes
numhhd1	race/gender recodes
nhhld1	race/gender recodes
alcage1	race/gender recodes
yrsc1c	race/gender recodes
alcwk1c	race/gender recodes
alc24hr1	race/gender recodes
highalc1	race/gender recodes