################################################################################################################

# Description and coding of variables in breast cancer dataset

# Dataset created by bootstrapping 20000 women from NCI-AARP dataset use in # Pfeiffer et al, PLos Medicine, 2013

# Author: Ruth Pfeiffer

# Date: April 23, 2018 ######################################################################

Abbreviations: HRT: hormone replacement therapy

EP: estrogen plus progestin HRT

BBD: benign breast disease

meno\_model = Age at Menopause, 0: miss 1:<=49 2: 50-54 3:55+

EPTDUR=EP or other HRT; 0: miss, 1:1-9y, 2: 10+y, 3: 10+y, 4: Other HRT

EPTDUR2= EP HRT; 0: none or unknown EP or Other HRT, 1: 1-9years, 2: 10+years

OtherHormones = Other HRT or Unk EP Indicator; 0:Otherwise, 1:Other HRT or Unknown type

EPTDUR\_missing = EP HRT and Other HRT missing Indicator; 0: used HRT 1: Missing HRT

fhbreast\_model= Family history of breast or ovarian cancer; 0: no, 1:yes

meno\_model= Age at menopause; 0:Pre or <49, 1: 50-54, 2: >=55

Age\_Meno = Age at menopause; 0:<49, 1:PreMen, 2:50-54, 3: >=55

PreMeno = Indicator of being premenopausal; 0:no, 1:yes;

BMI\_4\_cats: BMI< 25, 25 to < 30, 30 to < 35, and \_ 35kg/m2;

Parity\_0: 0: parous; 1: nulliparous

FC\_Age\_Trend: age at first birth: < 25, 25 − 29, and \_ 30 years;

bbd\_miss: missing Indicator for benign breast disease; 0: information available 1: Missing BBD

bbd\_model: benign breast diseases/biopsy: 0: no, 1: yes, 2: missing;

alc\_model: alcohol consumption: 1: 0,2: < 1, 3: >=1 drinks/day.

cig\_2 : being a former smoker: 0: no, 1: yes

cig\_3: being a current smoker: 0: no, 1: yes