1. Subject refused to answer or don’t know treated as missing

2. For Annual Family Income, subjects that answered over or under $20000 but not as a specific range are treated as missing

3. For Annual Family Income, create 4 categories: $0-$14,999; $15000-$34999; $35000-$64999; Over $65000.

4. For question ALQ120Q - How often drink alcohol over past 12 mos, people can answer 0 which means they never drink alcohol, and their corresponding information in ALQ130: Avg # alcoholic drinks/day -past 12 mos will be missing. We give these missing values a 0 in ALQ130. Categorize alcohol drinks per day into 3 categories: Non-drinker, <2 drinks per day, >=2 drinks per day.

5. Systolic Blood pressure have 4 readings; we use the average of these four readings. Categorize systolic blood pressure into fifth. These groups are: "Low", "Lower-middle","Middle","Upper-middle","High"

6. For depression, DPQ020, we define subject that answers several days or more than half the days or nearly every day as having depression.

7. History of cardiovascular disease is defined as participants has history of coronary heart disease or history of stroke, if one of them is missing, its value defined as the un-missing value, if both are missing, then its value is missing.

8. Menopausal status: if RHQ060 - Age at last menstrual period has a value, then define it as Postmenopausal, if RHQ020 - Age range at first menstrual period has a value and RHQ060 - Age at last menstrual period is missing, then define it as premenopausal. Categorize menopausal status as 2 groups: premenopausal, postmenopausal

9. We searched records of prescription medication use and identified users of aspirin, atovarstatin, ibuprofen, opium, statin, valsartan users. If we did not find any record for a medication for a person, then it is treated as non-user.

10. For total nutrition intake variables (TCAL, TCARB ….), we averaged day 1 and day 2. If one is missing, then the un-missing value is used. If both missing, then missing.

11. For dietary supplement use, if either day 1 or day 2 they used dietary supplement, then they are using dietary supplement overall. If neither day 1 and day 2 they used dietary supplement, they are not using dietary supplement. If one is missing, the un-missing value will be used. If both missing, then missing. For dietary supplement intake, categorize as Yes or No

12. If participants answered OCD150 - Type of work done last week to be not working at a job or business, their values in OCQ180 - Hours worked last week at all jobs will be missing. In this case, we give these missing values a 0 in OCQ180.

13. For occupation, we categorized participants into 3 categories: non-worker (0 hours a week), Part time worker (1-30 hours a week), Full time worker (>=31 hours a week)

14. For physical activity, if the participants either do a vigorous recreational activity or a moderate recreational activity, then they are defined as they do physical activity. If they neither do a vigorous recreational activity and a moderate recreational activity, then they are defined as they don’t do physical activity. If one is missing, the un-missing value will be used. If both missing, then missing. For physical activity, categorize it as 2 groups: "No activity","Vigorous or moderate activity"

15. People whose sedentary lifestyle minutes answered >=1000 is treated as missing because they are probably including the time they sleep and that is not plausible. Categorize sedentary lifestyle as fifth. For sedentary lifestyle, categorize as 5 groups: "Low", "Lower-middle", "Middle", "Upper-middle", "High"

16. People who answered they never smoked cigarettes regularly or not smoked at least 100 cigarettes in life or smoked at least 100 cigarettes in life but quit are categorized as non or light smoker. People who answered they smoked at least 100 cigarettes in life and continue smoking are categorized as follows: if >67 cigarettes (40 pack year) then heavily smoker, if <=67 cigarettes (40 pack year) then moderate smoker. For smoking, categorize as 3 groups: "Non or light smoker", "Moderate smoker", "Heavy smoker"

17. For individual food components, we averaged day 1 and day 2. If one is missing, then the un-missing value is used. If both missing, then missing.

18. Fruit consumption is defined as CITMLB+OTHER fruit, excluding juice.

19. Sea food consumption is defined as Seafood high in n-3 fatty acid + Seafood low in n-3 fatty acid.

20. Subject not eligible defined as ineligible for death linkage or age<18, not available for public release.

21. We consider age range 20-79

22. Missing lifestyle data: alcohol drinking, sleep, sedentary lifestyle, smoking

23. Missing personal and family history of disease: Missing history of hypercholesterolemia, history of hypertension, history of diabetes, history of depression, history of cardiovascular disease, history of cancer, family history of diabetes, family history of myocardial infraction

24. Missing demographic data: education, marital status, family annual income, PIR, occupation

25. Missing dietary variables: TKCAL, TCARB, TFIBE, TSFAT, TMFAT, TPFAT, TCHOL, TMAGN, F\_FRUIT, V\_TOTAL, PF\_SEAFD, G\_WHOLE, PF\_MPS\_TOTAL, PF\_MEAT PF\_CUREDMEAT, PF\_POULT, PF\_EGGS, PF\_NUTSDS, PF\_LEGUMES, D\_TOTAL D\_CHEESE and missing special diet and dietary supplement intake

26. Missing physical examination data: Systolic blood pressure, Missing BMI

27. Missing reproductive health information for women: missing Menopausal status, Hormone therapy, Parity, Oral contraceptive use

28. exclude Women are pregnant at baseline

29. exclude Implausible BMI (<15 or ≥60 kg/m2)

30. exclude Extreme value of total energy intake (<500kcal or >4500kcal)

31. For age, we create 3 age groups: 20-39 years old, 40-59 years old, 60-79 years old.

32. We categorize ratio of family income to poverty (PIR) into fifth to denote their socioeconomic status. 5 groups are : "Low", "Lower-middle","Middle","Upper-middle","High"

33. Participant are censored at last follow up 12.31.2019. Mortality data linked from CDC is up until 31 December 2019. We have 2007-2014 NHANES data.

34. Categorize BMI into 4 groups: Underweight (<18.5), Healthy weight (18.5<=<25), Overweight (25<=<30), Obesity (>=30), based on definition of CDC.

35. Borderline diabetes treated as no diabetes. History of diabetes as Yes or No.

36. Categorize sleep as 3 groups: <=4 hours, 5-8 hours, >=9 hours.

37. Convert ounce equivalent to grams for: cured meat (processed red meat), unprocessed red meat, poultry, seafood, whole grain, eggs, nuts and seeds, legumes. For total fruit, total vegetables, total diary, use cup equivalent.

38. Make unprocessed red meat as not only as continuous, but also quartiles and quintiles for standard model.

39. Divide unprocessed red meat by total energy and treat it not only as continuous, but also quartiles and quintiles for density model.

40. For gender, we create female and male group.

41. For education, we create 5 groups: "Less Than 9th Grade", "9-11th Grade (Includes 12th grade with no diploma)", "High School Grad/GED or Equivalent", "Some College or AA degree", "College Graduate or above".

41. For race, we create 5 groups: "Mexican American", "Other Hispanic", "Non-Hispanic White", "Non-Hispanic Black", "Other Race - Including Multi-Racial"

42. For marital status, we create 6 groups: "Never married", "Married", "Widowed", "Divorced", "Separated", "Living with partner"

43. For years of entering cohort, create 4 categories: "2007-2008", "2009-2010", "2011-2012", "2013-2014"

44. For history of diabetes, history of hypercholesterolemia, history of hypertension, history of depression, history of cardiovascular disease, history of cancer or malignancy, family history of diabetes, family history of myocardial infraction, category each of each as 2 groups: Yes or No

45. categorize hormone therapy use as Yes or No

46. categorize parity as two groups: Nulliparous and Parous

47. categorize oral contraceptive use as two groups: Yes or No

48. For use of aspirin, ibuprofen, opium, statin, valsartan, categorize two groups, Yes or No

49. For on special diet or not, categorize as Yes or No

50. For general health conditions, categorize as 5 groups: "Poor", "Excellent", "Very good", "Good", "Fair"

51. We have 48 optional adjusting variables in total. These variables are denoted by index (1-48). For each variable, drop a fair coin, if land 1, then include as adjusting variable, if land on 0, then not include as adjusting variable. We will have one group of adjusting variables. Do this to form n groups of adjusting variables. These n groups of adjusting variables can be seen as randomly sampled from 2^48 available groups of adjusting variables with replacement. Alcohol continuous and categorical could not be included at the same time. BMI continuous and categorical could not be included at the same time.

52. Cox-proportional hazards regression. HR and 95% CI, p value calculated.

53. For standard model: Treat unprocessed red meat consumption (gram) variable as continuous: 100 grams/day increase, or categorical (quintile or quartiles used: lowest as reference).

For multivariable nutrition density model: divide unprocessed red meat consumption (gram) by total energy intake, and then treat meat variable as continuous: 100 grams /2000kcal/day increase or categorical (quintile or quartiles used: lowest as reference).

54. The standard model (or so-called residual model, substitution model) that adjusted for total energy and the interpretation will be average relative causal effect.

55. The multivariable nutrition density methods that divided just the unprocessed red meat by the daily calorie intake. The interpretation will be average relative causal effect rescaled as a proportion of total energy. The density method also adjusted for total energy intake.

56. Subgroup analysis by age (4 groups), sex (3 groups).

57. Complete cases analysis.

58. 48 optional adjusting variables. 8 mandatory adjusted variables: For all model: age(continuous), sex, smoking, total energy are adjusted. For female model, menopausal status, hormone therapy, parity, oral contraceptive use are additionally adjusted.

59. We are only focusing on unprocessed red meat and all-cause mortality.

60. Accident constitutes all-cause mortality. All-cause mortality is defined as: diseases of heart, malignant neoplasms, chronic lower respiratory disease, accidents (unintentional injuries), cerebrovascular disease, Alzheimer disease, diabetes mellitus, influenza and pneumonia, nephritis nephrotic syndrome and nephrosis, all other causes (residual).