

The dataset final_menopause_EC has all new variables created.

For menopausal status there are 3 different variables created: menopausal status, reason for menopausal status and age at menopausal status

Menopausal status is given by the variable **menop_status**

Menop_status has the following codes

1="Premenopausal"

2="Postmenopausal"

9="Unknown menopausal status"

Reason for menopause is given by the variable **X_MENOPAUSE**. This variable has the following codes.

0 = "0=No information"

1 = "1=No menstruation, artificial menopause"

2 = "2=No menstruation, natural menopause"

3 = "3=No menstruation, hysterectomy"

4 = "4=No menstruation, IUD use"

5 = "5=No menstruation, other reason"

6 = "6=No menstruation, without any reason specified"

7 = "7=Menstruating while using an hormonal medication"

8 = "8=Premenopausal"

9 = "9=Menstruation currently regular"

10 = "10=Menstruation currently irregular"

11 = "11=Postmenopausal because of age"

12 = "12=corrected: post at baseline and pre at follow-up"

13 = "13=pre baseline and no 2nd questionnaire & < age at menopause"

14 = "14=pre baseline and unknow at 2nd questionnaire & < age at menopause"

15 = "15=unknown baseline and missing 2nd questionnaire & < age at menopause"

16 = "16=unknown / unknow & > age at menopause"

17 = "17= post / post differing age"

18 = "18= unknown / did not fill 2nd questionnaire"

19= "19= Hysterectomy - from patient registry"

20= "20= Unilateral oophorectomy - from patient registry"

21= "21= Bilateral oophorectomy - from patient registry"

And age at menopause is given by the variable: **X_MENOPAUSE_AGE_YRS_**

These 3 variables were created at 3 different points in time: at baseline, at follow-up and at end of follow-up. The philosophy behind being that future researchers might want to

ascertain exposures at different time, and will use the menopausal status variable that will fit their study design.

The variables for baseline are the variables names given above with the suffix_92_EC, for follow-up the suffix is _03_EC and _eof_EC at end of follow-up

How was the menopausal status variable built?

- A woman was considered to have an “unknown menopausal status” if she: - had a hysterectomy, started using HRT before end of period, or she did not give information on menopause
- A menopausal status at baseline and at follow-up (for women who filled the follow-up questionnaire was created using Agnes Fournier method)
- For women who did not filled the second questionnaire menopausal status at baseline was used to ascertain menopausal status at follow-up. For example, if she was postmenopausal at baseline then she was considered postmenopausal at follow-up. NOTE: some women, who filled both questionnaire, were postmenopausal at baseline and premenopausal at follow-up this women were changed to post-menopausal at follow-up
- For women with unknown menopausal status a age rule was used: if a woman was aged 55 or more (for never and former smokers) and 54 or + (for smokers) then she was considered postmenopausal
- For women with information on menopausal status and considered postmenopausal but with age at menopause then the age rule above was used to give an age at menopause
- The information from patient registry was used to complete a woman’s information on hysterectomy and unilateral or bilateral oophorectomy. This information was used to change a woman’s menopausal status if relevant. If a woman had a bilateral while she was considered pre menopausal, or with unknown menopausal status then her menopausal status was changed to post-menopausal. If a woman was considered post-menopausal but her aged at menopause was > than age at bilateral oophorectomy then her age at menopause was changed to age at surgery and the reason for menopause was also changed.

NOTE: 436 women gave differing date of menopause at baseline and follow-up (difference greater than 1 year). These women are listed in the SAS output