CHRIS Study

Interview – Gynecology

Version 1.1

24th April 2024

Authors: LB, MG

1. Introduction

This module stores information related to the menstruation cycle and the pregnancies history, that were collected at the interview, such as age at first menstruation, intake of contraceptives and its duration, number of pregnancies and their outcomes, and menopause occurrence.

Participants book a morning appointment at the CHRIS study center, ranging from 7.45 to 8.45 a.m. Each study participant is assigned a workflow at the reception. If there are ten study participants (maximum capacity), there are ten different workflows, marked with the letters from “A“ to “K“. The current workflow is as follows: A-B-C-D-E-F-G-H-I-K. All the workflows can be found in the documentation of CHRIS Baseline/General information/Administrative data, in the file named “Workflows at baseline assessment”. The interview occurs always after the spiralography and the blood drawing, for most as the last session, after the ECG assessment and the self-administered questionnaire (workflows B, C, E, F, H, I, L). For the remainder, the interview occurs after breakfast and just before the self-administered questionnaire (workflows A and G) or in between the blood drawing and the anthropometry (workflow D).

The interview full text and its corresponding answer lists are available at CHRIS Baseline/Interview.

2. History version changes

Version 1 of this interview module was in use between August 24th,2011 and November 2nd, 2012, whereas Version 2 had been in use between November 5th, 2012 and November 20th, 2013. Version 3 has been in use since November 21st, 2013.

Between the different versions, the following changes have occurred:

Version 1 to Version 2:

variables added: x0wo09, x0wo10, x0wo11, x0wo11a, x0wo12, x0wo12a

question order changed: x0wo03 (after x0wo02 => after x0wo01c)

question filtering criteria changed: x0wo02, x0wo03, x0wo04b, x0wo05

question rephrased: x0wo09 (it), x0wo10 (it)

response options changed x0wo02 (added "3 I don't know"), x0wo03 ("1 Yes" instead categories 1 to 5), x0wo06 (added "3 I don't know")

Version 2 to Version 3:

variables dropped: x0wo01c

variables added: x0wo01

question order changed: x0wo01a, x0wo01b (both if x0wo01=”Yes”)

The rephrasing of some questions and answers has occurred only in its Italian version.

3. Data cleaning

1. The main CHRIS dataset was loaded.
2. The variables with further comments were all translated, x0won1, x0won2, x0won3, x0won4, and x0wonote, and categorized when possible.
3. All variables had their missing observations set to “Missing by design” when the participant was a male.
4. It was observed that those females, for which the sex variable x0_sex had wrongly been assigned, could not complete this module and have therefore no information available.
5. The question on the menarche (first menstruation) could either be answered with the year (x0wo01a) or the age (x0wo01b), according to what the participant recalled best. The answers have been standardized into age (x0wo01b).
6. Then the observation of x0wo01b were transformed as follows:
    a) into “Don’t know” (-88), if x0wo01b was 99 or 9999;
    b) into “Missing by design” (-99) if x0wo01b was missing and the participant was male;
    c) into “Don’t know” if x0wo01b was missing and x0wo01c==”No”;
    d) into “Unexpected missing” (-89)
7. The observations of the dichotomous variable on occurrence of the first menstruation, x0wo01, added in the third version, were transformed as follows:
    a) The answers “Yes” and “No” were assigned to the 1 and 2 values;
    b) The answer “Don’t know” was assigned to the missing type -88;
    c) The variable was transformed into “Yes” if the age of first menstruation, x0wo01b, was given;
    d) The missing observations were transformed to “Don’t know” if the age at first menstruation was not known;
    e) The remaining missing values were transformed into “Unexpected missing” (-89).
8. The variables of “year of first menstruation”, x0wo01a, and “No menstruation yet”, x0wo01c, were dropped.
9. The answer options to the current contraceptive intake question, x0wo03, were harmonized between the different versions, into “Yes”, “No”, and “Refuse to answer” (-87) options. The missing observations of x0wo03 are transformed as follows:
    a) “Missing by design” (-99) if the interview version, x0wover, is the first and the participant never took contraceptives (x0wo02=”No”);
    b) “Missing by design” (-99) if the interview version, x0wover, is after the first and the participant age is above 60;
    c) “Unexpected missing” (-89) otherwise.
10. The variable on contraceptive intake ever, x0wo02, had its missing observations transformed into:
    a) “Missing by design” (-99) if the interview version, x0wover, was the second and the current contraceptive intake is either “Yes” or “Don’t know”;
    b) “Missing by design” (-99) if the interview version, x0wover, was the third and the current contraceptive intake is either “Yes” or “Don’t know” or the age is above 60;

c) "Unexpected missing" otherwise.

11. The missing observations on contraceptives intake duration in years, x0wo04b, were transformed as follows:
    a) Into "Missing by design" if the participant never took contraceptives (x0wo02="No") or could not answer to that question (x0wo02="Missing by design");
    b) The number of months of intake duration transformed into years, x0wo04/12, where x0wo04a was available and x0wo04b not;
    c) "Unexpected missing" otherwise.
12. The ever contraceptives intake, x0wo02, was corrected into "Yes" if it was missing but a positive contraceptives intake duration, x0wo04b, was provided.
13. The variable of contraceptives intake duration in months, x0wo04a, was dropped.
14. The variable on current pregnancy status, x0wo05, had its missing observations transformed into:
    a) "Missing by design" if the age of the participant, x0_age, is above 55;
    b) "Unexpected missing" otherwise
15. The variable on the week of the current pregnancy, x0wo05a, had its missing observations transformed into:
    a) "Missing by design" if the current pregnancy status, x0wo05, is either negative or unknown or "Missing by design";
    b) "Unexpected missing" otherwise
16. The variable on ever being pregnant, x0wo12, had its missing observations transformed into:
    a) "Yes" if the number of born children and stillbirths, x0wo12a, was positive;
    b) "Not in use" if the interview version was the first;
    c) "Unexpected missing" otherwise.
17. The variable on the number of pregnancies with a live birth or stillbirth (birth of dead fetus after at least 28 weeks), x0wo12a, had its missing observations transformed into:
    a) "Not in use" if the interview version was the first;
    b) "Missing by design" if the question on pregnancy ever, x0wo12, was "No" or "Missing by design";
    c) "Unexpected missing" otherwise.
18. Inconsistencies between number of completed pregnancies, x0wo12a, and number of born children, x0pe11, was checked.
19. The variable on current regular menstruation, x0wo06, had its missing values transformed into "Unexpected missing" and its third answer option into -88 ("Don't know").
20. The variable on reasons for stop of the menstruation, x0wo08a, was translated and categorized.
21. For those participants reporting pregnancy, lactation or contraceptives as reasons for definitive interruption of menstruation, the answer to regular menstrual bleedings, x0wo06, was changed from "No" to "Yes", and their reason for menstruation stop was changed to "Missing by design".
22. For those participants reporting the tubes ligation or another surgical operation as cause for their menopause, their reason for the menstruation stop was changed to "Operation".
23. For those participants reporting hormone replacement therapy as a reason for menstruation stop in x0wo08a, their answer to x0wo10 (hormone replacement therapy ever) was changed from missing to "Yes".

24. The variables on hormone replacement therapy, x0wo09, x0wo10, and x0wo11, had their missing observations transformed into “Not in use” if the interview version, x0wover, was the first.
25. The dichotomous variable on current hormone replacement therapy intake, x0wo09, had its missing observations turned into:
    a) “Missing by design” (-99) if the reason for stopped menstruation, x0wo08, is an operation or “Another reason” or “Missing by design”
    b) “Unexpected missing” otherwise
26. The dichotomous variable on current hormone replacement therapy intake, x0wo09, had its third answer option turned into “Don’t know” (-88).
27. The dichotomous variable on ever hormone replacement therapy intake, x0wo10, had its missing observations turned into:
    a) “Missing by design” (-99) if the current intake variable, x0wo09, is either “Yes” or “Missing by design”
    b) “Unexpected missing” otherwise
28. The duration of the hormone replacement therapy could be given in years, x0wo11a, or in months, x0wo11b. The variable of therapy duration in years, x0wo11a, was replaced with x0wo11b/12 if it was missing and the duration in months was instead available.
29. The missing observations of x0wo11a were transformed into:
    a) “Missing by design” if both current and ever hormone replacement therapy intake were “Missing by design”;
    b) “Missing by design” if both current and ever hormone replacement therapy intake were “No”;
    c) “Unexpected missing” otherwise;
    d) If its values were 99, then it was replaced with “Don’t know”.
30. The variable on hormone replacement therapy duration in months, x0wo11b, was dropped.
31. The baseline dataset was saved.

4. Advices for the analysis

The analyst should read the notes saved in the variables x0won1-x0won4 and x0wonote, where multiple contraceptives, stillbirths, and miscarriages, as well as hormone replacement therapies are reported.

The free text question on reasons for menstruation stop, x0wo08a, has been filled also with information that are not cause for menstruation stop, but rather medication to treat menopause symptoms. Similarly, other participants reported not to have regular bleeding because they were taking contraceptives, they were pregnant or lactating: in those cases, when it was clear the menstruation stop was temporary, their answer to the menstruation stop was changed.

Finally, the analyst should always take into account that the operator in charge of carrying out the interview might have influenced how the participant reported their answers. The analyst should therefore adjust for the operator variable, x0_opintc, when possible.