CHRIS Study

Interview – Kidney Diseases

Version 1.1

24th April 2024

Authors: LB, MG

1. Introduction

This module stores information related to the diseases of the kidney of the participant, that were collected at the interview.

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. This module is based on the Renal DataSchema survey, developed by the multidisciplinary consortium of experts DataSHaPER (DataSchema and Harmonization Platform for Epidemiological Research; http://www.datashaper.org), pulled together and coordinated by three international organizations: P^3G (Public Population Project in Genomics), PHOEBE (Promoting Harmonization of Epidemiological Biobanks in Europe) and CPT (Canadian Partnership for Tomorrow Project).

2. History version changes

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

Version 1 to Version 2

variables dropped: x0ki01c, x0ki02c, x0ki02d, x0ki04c, x0ki05c, x0ki06, x0ki06b, x0ki06c, x0ki06d, x0ki07c, x0ki08c, x0ki09d, x0ki16, x0ki17, x0ki17b, x0ki18, x0ki18b

variables added: x0ki21c, x0ki22, x0ki22b, x0ki23, x0ki23b

question order changed: x0ki08d (before: after x0ki08c, then: after x0ki08)

question filtering criteria changed: x0ki10 (only if x0ki00=1 or x0ki09=1), x0ki19 (no restriction, before if x0ki16=1), x0ki20 (no restriction, before if x0ki16=1), x0ki21 (no restriction, before if x0ki16=1)

question rephrased: x0ki01 (de, it), x0ki01a (de, it), x0ki01d (de, it), x0ki02 (de, it), x0ki02a (de, it), x0ki04 (it), x0ki04a (de, it), x0ki04d (de, it), x0ki05 (it), x0ki05d (de, it), x0ki08 (it), x0ki08 (de, it), x0ki19 (de,it)

other: Vasculitis part (x0cd14, x0cd14b, x0cd14c, x0cd14d) was moved to the Chronic diseases module

Version 2 to Version 3

question filtering criteria changed: x0ki10 (only if x0ki09=1), x0ki22 (only if x0ki09=1)

3. Data cleaning

1. The main CHRIS dataset was loaded.
2. The variables on kidney diseases ever diagnosed, x0ki00, had it observations transformed into:
    a) "Unexpected missing" (-89) if they were missing,
    b) "Don't know" (-88) if the third answer option "I do not know" was chosen.
3. The variables on diagnosis of glomerulonephritis, pyelonephritis, renal arteries disease, congenital kidney disease, kidney stones, and other kidney diseases, x0ki01, x0ki02, x0ki04, x0ki05, x0ki07, and x0ki08, had their observations transformed into:
    a) "Missing by design" (-99) if they were missing and no kidney disease ever was reported (x0ki00="No", "Missing by design" or "Don't know")
    b) "Unexpected missing" if they were still missing,
    c) "Don't know" (-88) if the third option "I do not know" was chosen.
4. The variables on the age at diagnosis of glomerulonephritis, pyelonephritis, renal arteries disease, congenital kidney disease, kidney stones, and other kidney diseases, x0ki01b, x0ki02b, x0ki04b, x0ki05b, x0ki07b, and x0ki08b, had their observations set to:
    a) "Missing by design" (-99) if they were missing and that specific disease was not reported (x0ki0*= "No", "Missing by design", "Not in use" or "Don't know"),
    b) "Unexpected missing" if they were missing,
    c) "Don't know" if the age reported was 99.
5. The variables on treatment of glomerulonephritis, pyelonephritis, renal arteries disease, congenital kidney disease, kidney stones, and other kidney diseases, x0ki01c, x0ki02c, x0ki04c, x0ki05c, x0ki07c, and x0ki08c, had their missing observations set to:
    a) "Not in use" (-98) if the version of the questionnaire x0kiver was not the first,
    b) "Missing by design" (-99) if that specific disease was not reported (x0ki0*= "No", "Missing by design", "Not in use" or "Don't know"),
    c) "Unexpected missing" otherwise.
6. The variable on diagnosis of kidney cancer, x0ki06, had its missing observations set to:
    a) "Not in use" if the version of the questionnaire x0kiver was not the first,
    b) "Missing by design" if no kidney disease ever was reported (x0ki00="No", "Missing by design", "Not in use" or "Don't know")
    c) "Unexpected missing" otherwise.
7. The variables on age at diagnosis and treatment of kidney cancer, x0ki06b and x0ki06c, had their missing observations set to:
    a) "Not in use" if the version of the questionnaire x0kiver was not the first,
    b) "Missing by design" if no kidney cancer was reported (x0ki06="No", "Missing by design" or "Don't know")
    c) "Unexpected missing" otherwise.
8. If in the variable on the other kidney diseases, x0ki08, a pyelonephritis ("Nierenbeckenentzündung" in German) was reported, the corresponding variable x0ki02 was corrected accordingly and x0ki08 was set to "Missing by design". Similarly, if a congenital

condition was reported, it was assigned to x0ki05; if glomerulonephritis was reported it was assigned to x0ki01.

9. All the reports of kidney cancer in the cancer module (x0ca01, x0ca02, x0ca03, x0ca04, x0ca05) of the interview were reported also in x0ki08, to make sure all kidney related information were contained also in the kidney module.
10. All the reports of glomerulonephritis, pyelonephritis, or kidney sand in the x0cd14d variable were also assigned to the corresponding variables, x0ki01, x0ki02, and x0ki08, respectively.
11. The variable on diagnosis of kidney failure, x0ki09, had its observations transformed into:
    a) "Missing by design" (-99) if they were missing and no kidney disease ever was reported (x0ki00="No", "Missing by design" or "Don't know")
    b) "Unexpected missing" if they were still missing,
    c) "Don't know" (-88) if the third option "I do not know" was chosen.
12. The variable on the age at diagnosis of kidney failure, x0ki09b, had its observations set to:
    a) "Missing by design" (-99) if they were missing and kidney failure was not reported (x0ki09= "No", "Missing by design" or "Don't know"),
    b) "Unexpected missing" if they were missing,
    c) "Don't know" if the age reported was 99.
13. The variable on current kidney failure, x0ki09c, had its observations classified into yes and no when the free text contained a "Ja" (Yes) or a "Nein" (No), and the remaining observations set to:
    a) "Missing by design" (-99) if they were missing and kidney failure was not reported (x0ki09= "No", "Missing by design" or "Don't know"),
    b) "Unexpected missing" (-89) if they were missing,
    c) "Don't know" (-88) if the chosen option was "I do not know".
14. The variable on kidney failure treatment, x0ki09d, had its missing observations set to:
    a) "Not in use" (-98) if the questionnaire version, x0kiver, was not the first,
    b) "Missing by design" (-99) if kidney failure was not reported (x0ki09= "No", "Missing by design" or "Don't know"),
    c) "Unexpected missing" (-89) otherwise.
15. The variable on kidney transplantation, x0ki10, had its observations transformed into:
    a) "Missing by design" (-99) if they were missing, the questionnaire version was the second, no kidney disease ever was reported (x0ki00="No", "Missing by design" or "Don't know"), and kidney failure was not reported (x0ki09="No", "Missing by design" or "Don't know"),
    b) "Missing by design" (-99) if they were missing, the questionnaire version was the third, no kidney disease ever was reported (x0ki00="No", "Missing by design" or "Don't know"), and kidney failure was not reported (x0ki09="No", "Missing by design" or "Don't know"),
    c) "Unexpected missing" if they were still missing,
    d) "Don't know" (-88) if the third option "I do not know" was chosen.
16. The variable on the number of kidney transplantations, x0ki10a, had its missing observations set to:
    a) "Missing by design" (-99) if kidney transplantation was not reported (x0ki10= "No", "Missing by design" or "Don't know"),

b) “Unexpected missing” (-89) otherwise.

17. The variables on the age at each kidney transplantation, x0ki11b, x0ki12b, x0ki13b, x0ki14b, and x0ki15b, had their missing observations set to:
    a) “Missing by design” (-99) if the number of kidney transplantations reported in x0ki10a was positive but less than their position (e.g., for x0ki12b, if x0ki10a>0 and x0ki10a<2) or if no kidney transplantation was reported (x0ki10= “No”, “Missing by design” or “Don’t know”),
    b) “Unexpected missing” (-89) otherwise.
18. The variable on other kidney surgeries, x0ki16, had its observations set to:
    a) “Not in use” (-98) if they were missing and the questionnaire version, x0kiver, was not the first,
    b) “Unexpected missing” (-89) if they were still missing,
    c) “Don’t know” (-88) if the option “I do not know” was chosen.
19. The variables on kidney stone surgery and on surgery for other kidney problems, x0ki17 and x0ki18, had their missing observations set to:
    a) “Not in use” (-98) if the questionnaire version, x0kiver, was not the first,
    b) “Missing by design” (-99) if a kidney surgery was not reported (x0ki16= “No”, “Missing by design” or “Don’t know”),
    c) “Unexpected missing” (-89) otherwise.
20. The variables on the age at the kidney stone surgery and on the age at the surgery for other kidney problems, x0ki17b and x0ki18b, had their missing observations set to:
    a) “Not in use” (-98) if the questionnaire version, x0kiver, was not the first,
    b) “Missing by design” (-99) if that specific surgery was not reported (e.g., for x0ki17b, if x0ki17= “No”, “Missing by design” or “Don’t know”),
    c) “Unexpected missing” (-89) if that specific surgery was reported or “Unexpected missing”,
    d) “Don’t know” if the reported age was 99 and that specific surgery was reported or “Unexpected missing”.
21. The variables on kidney donation, angioplasty of renal arteries, and further kidney surgeries, x0ki19, x0ki20, and x0ki21, had their missing observations set to:
    a) “Missing by design” (-99) if a kidney surgery was not reported (x0ki16= “No”, “Missing by design” or “Don’t know”) and the version was the first,
    b) “Unexpected missing” (-89) otherwise.
22. The variables on the age at kidney donation, angioplasty of renal arteries, and further kidney surgeries, x0ki19b, x0ki20b, and x0ki21b, had their missing observations set to:
    a) “Missing by design” (-99) if that specific surgery was not reported (e.g., for x0ki19b, if x0ki19= “No”, “Missing by design” or “Don’t know”),
    b) “Unexpected missing” (-89) otherwise.
23. The variable on dialysis, x0ki22, had its missing observations set to:
    a) “Not in use” if the questionnaire version was not the first,
    b) “Missing by design” (-99) if the questionnaire version was the second, no kidney disease ever was reported (x0ki00=”No”, “Missing by design” or “Don’t know”), and kidney failure was not reported (x0ki09=”No”, “Missing by design” or “Don’t know”),

c) “Missing by design” (-99) if the questionnaire version was the third, no kidney disease ever was reported (x0ki00=”No”, “Missing by design” or “Don’t know”), and kidney failure was not reported (x0ki09=”No”, “Missing by design” or “Don’t know”),
d) “Unexpected missing” (-89) otherwise.

24. The variables on age at dialysis start and on current dialysis, x0ki22b and x0ki23, had their missing observations set to:
    a) “Not in use” if the questionnaire version was not the first,
    b) “Missing by design” (-99) if no dialysis was reported (x0ki22=”No”, “Missing by design” or “Don’t know”),
    c) “Unexpected missing” (-89) otherwise.
25. The variables on age at dialysis interruption, x0ki23b, had its missing observations set to:
    a) “Not in use” if the questionnaire version was not the first,
    b) “Missing by design” (-99) if no dialysis was reported (x0ki23=”Yes” or “Missing by design”),
    c) “Unexpected missing” (-89) otherwise.
26. All the year variables, x0ki01a, x0ki02a, x0ki04a, x0ki05a, x0ki06a, x0ki07a, x0ki08a, x0ki09a, x0ki11a, x0ki12a, x0ki13a, x0ki14a, x0ki15a, x0ki17a, x0ki18a, x0ki19a, x0ki20a, x0ki21a, x0ki22a, and x0ki23a, were dropped in favor of the age variables, x0ki01b, x0ki02b, x0ki04b, x0ki05b, x0ki06b, x0ki07b, x0ki08b, x0ki09b, x0ki11b, x0ki12b, x0ki13b, x0ki14b, x0ki15b, x0ki17b, x0ki18b, x0ki19b, x0ki20b, x0ki21b, x0ki22b, and x0ki23b.
27. The variables specifying the diagnoses or surgery type, x0ki01d, x0ki02d, x0ki04d, x0ki05d, x0ki06d, x0ki08d, and x0ki21c, were translated and categorized when possible.
28. The variables storing the notes additional information on kidney diseases, kidney surgery, and kidney therapy, x0kin1, x0kin2, x0kin3, and x0kinote, were translated and categorized when possible.
29. The baseline dataset was saved.

4. Advices for the analysis

The content of the nurse’s notes, referring to kidney diseases, can include information on multiple kidney diagnoses, as well as multiple kidney surgeries. Additionally, circumstances of kidney malfunction (e.g. after a certain treatment) or borderline values are reported in x0kinote.

The current medications can be looked at in the drugs module x0dd*, where the participant let their current medication packages be scanned by the nurse at the study center.

Due to the availability in German of common words to describe most conditions, the common word for glomerulonephritis is “Nierenentzündung”, the one for pyelonephritis is “Nierenbeckenentzündung”, and the ones for renal failure are “Nierenversagen” or “Niereninsuffizienz”. Consequently, many participants were unaware of their condition’s correct medical term and it the interviewing operator had to correctly assign the kidney disease and to write down the disease denomination. Therefore, from the second version the terms Nierenentzündung and Nierenbeckenentzündung were added in the questions to better specify them. Additionally, the analyst can use the blood values for creatinine (x0lp07), the urine values for creatinine (x0lp01), albumin (x0lp64), and the composite biomarkers such

as urine albumin creatinine ratio UACR (x0lp63) and estimated Glomerular filtration rate eGFR (x0lp86a and x0lp86b) to assess the renal failure stage.

The self-reported current presence of renal failure at the time of the baseline assessment, stored in x0ki09c, has multiple missing observations among those participants reporting renal failure at some point of their life. Consequently, to better characterize current renal failure, the analyst should also consider the blood and urine values to complement this information, that were already mentioned in the previous paragraph.

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.

5. References

Renal Schema: Fortier I, Burton P R, Robson P J, et al. Quality, quantity and harmony: the DataSHaPER approach to integrating data across bioclinical studies, Int J Epidemiol, 2010 Oct; 39(5): 1383–1393. DOI: 10.1093/ije/dyq139