CHRIS Study

Lab routine measurements

Version 1.1

24th April 2024

Authors: LB, MG

1. Introduction

Blood and urine samples (together referred to as biochemical traits) were collected for all participants at the CHRIS study center in Schlanders hospital.

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”.

When making the appointment, participants are asked to stay fasting overnight and to drink only water before coming to the study center. The precise set of instructions are available at CHRIS Baseline/General Information in the files named “Confirmation letter – German version” and “Confirmation letter – Italian version”. During the self-administered questionnaire session, the participant fills in the blood collection section where they can report deviations from the instructions given. Participants can have a breakfast at the study center after the collection of the biological samples.

Urine containers are labelled by the study assistant responsible for the blood drawing and marked with the corresponding letter of the workflow.

The participant is given the corresponding empty urine collection container. The urine collection container must be filled before breakfast (at least 30 ml of urine) and handed over at reception immediately after filling, as it has to be placed in the refrigerator in the blood collection room as soon as possible.

The participant is also given a blood withdrawal form on paper, where a limited set of questions on anticoagulants intake, hepatitis, hemophilia, and problems by blood withdrawal are asked. Before taking the blood sample, the nurse checks the completed blood collection form (CHRIS Baseline/Biochemical traits, in the files named “Blood collection form – German version”, “Blood collection form – Italian version”, “Blood collection form – English translation”). The nurse asks the participant, who has taken off their jacket/pullover, to sit comfortably on the blood collection chair and to uncover the veins of the arm’s crook. Immediately before the blood drawing, the study assistant compares the participant number on the workflow and on the test tubes, then checks the date of birth, removes the label (collection date) and sticks it on the sheet prepared for this purpose (This is the confirmation that the study assistant has checked the participant's year of birth and that it is correct). More details on the blood intake are available in the file “Blood draw and sample transport SOP” in CHRIS Baseline/Biochemical traits.

Given the instability of some of the blood parameters, pre-analytical sample processing is performed immediately after blood drawing at the study center by the responsible nurse, to ensure reliability and accuracy of all the measurements.

After the pre-analytical sample processing, the blood and urine samples are transported daily at 11.00 a.m. from Schlanders (CHRIS center) to Merano (Franz Tappeiner hospital) by routine hospital transport (45-75 min):

1. test tubes with blood for the determination of the selected blood parameters (4°C/room temperature (RT));

2. urine collection containers for the determination of the selected urine parameters (4°C) plus two empty test tubes (with labels).

The samples were then processed in the EURAC laboratory CHRIS/routine laboratory in the Merano hospital as follows:

1. Five test tubes for blood and two test tubes urine analyses (routine laboratory hospital Merano for the determination of blood and urine parameters.

2. Six blood test tubes for the biobank (EURAC Laboratory CHRIS Hospital Merano)

Additional information on the procedures in place to ensure a correct preservation of the samples until their arrival at the laboratory in Merano hospital and to the biobank facilities is available in the files "Blood draw and sample transport SOP" and "Blood draw tubes labels" in CHRIS Baseline/Biochemical traits.

The laboratory machines used to obtain the blood and urine parameters are listed in table 1 and table 2, respectively, of this document. Access to the manuals will be granted only after a data/material transfer agreement is in place, due to copyright issues.

2. History of the laboratory changes

During the baseline collection, from 2011 to 2018, the laboratory of Merano hospital underwent some machinery updates, that were reflected in the measurement of biochemical traits of the CHRIS study participants. Furthermore, when technical issues arose, test tubes needed be sent to another laboratory, based at Bolzano hospital. In this section the main changes that do not refer to machine changes are reported.

2.1 Blood samples

Antithrombin, x0lp04: new kit/machine (Stago) since January 30th, 2014. Plasma: samples taken between January 14th, 2014 and January 24th, 2014 were stored at -30°C until the arrival of the new kit/machine.

Prothrombin time (sec), x0lp02d: measured since January 30th, 2014.

Transferrin saturation, xolp31: since June 8th, 2018, its values are missing. The values are calculated internally from iron and transferrin.

Homocysteine, x0lp32: between December 3rd, 2015 and March 15th, 2016 due to technical problems at the Merano hospital lab the measurement is performed at the laboratory of the Bolzano hospital (machine: Roche MODULAR ANALYTICS)

Homocysteine, x0lp32: data collection finished on November 17th, 2017.

2.2 Urine samples

The urine parameters underwent no specific changes. Machine updates as reported in table 2.

3. Data cleaning

Here the procedure of cleaning of the biochemical traits related variables is described.

1. The dataset with all baseline assessments (interview, self-administered questionnaire, anthropometry, biochemical traits) was loaded. The biochemical traits variables all start with x0lp.
2. Variables which, outside the CHRIS protocol, were measured for clinical reasons for one participant were erased.
3. The values of homocysteine, as they come from two different labs depending on the measurement day, are stored in two separate variables, x0lp32 (Merano lab) and x0lp32a (Bolzano lab). They were merged into x0lp32.
4. The information on antinuclear antibodies was also stored in two variables and then merged into unique ones, namely x0lp51a, x0lp51b, x0lp51c and x0lp51h.
5. For each variable, reasons for missingness have been translated and standardized and stored in its own note variable, with an n at the end of its name (x0lp01n is the note variable of x0lp01). Examples of reasons reported are "sample not handed in", “sample not received”, “not sufficient material”, “test tube arrived broken”, “interference by lipemia”, “interference by icterus”, “interference by hemolysis”, “test tube broken during centrifuge”, “sample not suitable”.
6. When detection limits were present, the observations were replaced with the closest value into a new variable ending with q, taking into account the precision of that variable. For instance, C-reactive protein (x0lp33) was undetectable below 0.02. When observations were “<0.02”, the variable x0lp33q was assigned 0.01. Antithrombin (x0lp04) could not be detected above 160, so its observations “>160.0” were changed in x0lp04q into 160.1
7. For each variable, its unit variable was created, to enable transformations in case a variable had had changes in the unit. This was the case of Calcium (x0lp23), Phosphor (x0lp25), and Magnesium (x0lp26), all measured before in mg/dL and later in mmol/L. The variables with all observations in mg/dL are x0lp23, x0lp25, x0lp26, those in mmol/L are x0lp23a, x0lp25a, x0lp26a.
8. The measured Calcium in mmol/L (x0lp23a) was corrected by albumin according to this formula, when the second machine (fill IN HERE) was in place:
   Calcium corrected [mmol/L] = Calcium measured [mmol/L] - 0.25 x Albumin [g/dl] + 1,
   storing its values in the variable x0lp24. This value was not automatically calculated after calcium unit change (27/11/2012).
9. Consistency between Iron (x0lp27) and Total iron-binding capacity, directly proportional to Transferrin (x0lp29), was checked.
10. Transferrin saturation (x0lp31) was calculated as
    transferrin saturation=round((Iron [µg/dl]/Transferrin [mg/dl])*70.783166,.1)
11. The formulas of relationship among red blood cells were checked. MCV=mean corpuscular volume, MCH=mean corpuscular hemoglobin, MCHC= mean corpuscular hemoglobin concentration, RBC=red blood cells.

MCV (fL) = Hematocrit (in %) × 10 / RBC (in 10-12/L) [1]
MCH (pg) = hemoglobin (in g/dL) × 10 / RBC (in 1012/L)[2]
MCHC (g/dL) = hemoglobin (in g/dL) × 100 / Hematocrit (in %)[3]
MCHC (g/dL) = MCH (in pg) × 100 / MCV (in fL)

12. The variable thyroid stimulating hormone (TSH) was merged from three separate variables into x0lp35.
13. F T3 and F T4 measured only for a subset with high or low TSH, namely until April 14th, 2014, those at least 50 years old and with TSH<0.2 or ≥3.8, those below 50 years old with TSH<0.3 or TSH≥4.5, and, since April 15th, 2014, those older than 19 years old with TSH<0.4 or TSH≥4.0, and those at most 19 years old and TSH <0.5 or TSH>4.3. For those who were not fulfilling the criteria described above, the missing values were replaced with "missing by design" (-99).
14. The variable prothrombin time (x0oc02d) had its missing values replaced with "not in use" (-98) if the examination date was before January 30th, 2014.
15. Similarly, the interference indexes (lipemic, hemolysis, icterus), x0oc82, x0oc83, x0oc84, that were measured starting from April 8th, 2014, had its missing values set to "Not in use" (-9998) when the examination date was before that date
16. The HIL ABBOTT variable, x0oc85, in use since April 8th, 2014, had its missing values set to "Not in use" (-98) before April 8th, 2014.
17. The parameter homocysteine, measured until November 17th, 2017, had its missing values replaced with "Not in use" (-98) after that date.
18. All the quantitative variables have then the remaining missing values set to "unexpected missing" (-89 or -9989, according to whether they include negative values).
19. The variables x0lp51a and x0lp51f on antinuclear antibodies (ANA) had a mixture of text and numeric values, its observations with value "negative" and "positive" have been replaced with 0 and 1, respectively.
20. If x0lp51a=0, then x0lp51f was not measured and therefore set as "missing by design" (-99).
21. The titre used to observe the ANA (x0lp51c, x0lp51e, x0lp51g) was ordered in descending order, from 1:80 to 1:20480.
22. The titre and ANA patterns variables (x0lp51b-x0lp51e, x0lp51g) had their missing values set to:
    a) "missing by design", if x0lp51a=0
    b) "unexpected missing" if x0lp51a=1 or x0lp51a="unexpected missing"
23. The observations of the ANA pattern stored in x0lp51b and x0lp51d was standardized into 19 and six categories, respectively.
24. The urine color variable, x0lp61b, was returned by the lab machine. After a descriptive analysis of its distribution, it can be observed that its frequencies are not stable over time and between languages, it is not considered of sufficient quality to be shared.
25. The urine appearance variable, x0lp61c, has been standardized into three categories: clear, slightly cloud, and cloudy.
26. For each variable, its correspondent machine changes were stored in a variable with an "m" at the end of the variable name. For instance, x0lp32m stores the machines used to measure the parameter x0lp32, already considering the machine upgrading underwent by the laboratory.

[1] https://www.labpedia.net/mean-corpuscular-volume-mcv-mean-cell-volume/
[2] https://www.labpedia.net/mean-corpuscular-hemoglobin-mch-mean-cell-hemoglobin/
[3] https://www.labpedia.net/red-blood-cell-rbc-part-4-blood-indices/

27. Possible duplicates between participants on multiple parameters were searched for and only one case with two women with the same values on around 10 parameters arose. The values were attributed to the participant more likely to have those values, i.e. the younger one. The other participant had its probably duplicate values set to missing.
28. The parameters estimated Glomerular filtration rate (eGFR) CKDEPI crea and MDRD4 were computed using the functions CKDEpi.creat and MDRD4, of the R package nephro. Its results were stored and merged back to the baseline dataset using the AID of the participants. The variable names changed into x0lp86a and x0lp86b, respectively.

4. Advices for the analysis

When analyzing any parameters here described, the machine changes that the laboratory in Merano underwent must be taken into account. Indeed, even after adjusting for age and sex, the distributions of some parameters differ substantially between the two measurement methods.

In order cases, the unit of measurement was changed over time, resulting in a change of precision over time. This is the case for variables like x0lp23-x0lp26, that were first reported in mg/dL and then in mmol/L.

When analyzing the blood and urine parameters, the following information must be considered.

1. The hemolysis, icterus, and lipemia can interfere with biochemical traits. This information is available in x0lpnote and by looking at the x0lp82-x0lp84 variables.
2. the fasting status of the participants. This information was collected in the blood collection form, saved in the variables with name x0bc*.
3. specific drugs might have interference. Information on drugs is available in the variables with name x0bl02*, x0bl12a, x0pk11, x0hc34, and the whole drug summary data x0dd*.

The analyst might consider excluding those observations where interference might have occurred.

Further information on the interferences occurring at the level of each blood or urine parameter is stored in the file “Lab routine measurements' interferences list”.

5. References

Noce D, Gögele M, Schwienbacher C, Caprioli G, De Grandi A, Foco L, Platzgummer S, Pramstaller PP, Pattaro C. Sequential recruitment of study participants may inflate genetic heritability estimates. *Hum Genet*. 2017 Jun;**136**(6):743-757. DOI: 10.1007/s00439-017-1785-8.

nephro R package: Pattaro C, Riegler P, Stifter G, Modenese M, Minelli C, Pramstaller PP. Estimating the glomerular filtration rate in the general population using different equations: effects on classification and association. *Nephron Clin Pract* 2013; **123**(1-2):102-11. DOI: 10.1159/000351043

Nephro R package overview: https://rdrr.io/cran/nephro/man/nephro-package.html

Nephro R package on CRAN: https://CRAN.R-project.org/package=nephro