

# **CHRIS Covid-19 Study**

## **Questionnaire – Baseline information**

**Version 1.0**

**9 May 2022**

Authors: MG

## 1. Introduction

This module stores information related to self-reported sociodemographic characteristics, lifestyle, health status, and anthropometry.

The CHRIS COVID-19 study was designed to estimate the distribution of SARS-CoV-2 infection cases in Val Venosta/Vinschgau since 1 February 2020, as well as the proportion of asymptomatic individuals among positive cases, to characterize transmission within households, to assess the relationship between antibody response and disease severity, to observe the evolution of antibody response over time, to identify environmental, molecular and genetic risk factors, to identify long-term sequelae.

The CHRIS COVID-19 study was organized in three stages:

*Stage 1:* A stratified random sample of 1812 CHRIS study participants was selected to represent the adult population of the Venosta/Vinschgau district. Out of this sample, 845 CHRIS participants replied to an online or paper questionnaire, underwent a molecular test based on a nasopharyngeal swab and a serum antibody test.

*Stage 2:* All 13,393 CHRIS study participants and their consenting cohabitants were invited to fill in an online questionnaire on their past and current health status, and on SARS-CoV-2 (potential) exposure and testing. Each CHRIS participant received ten access tokens to let them and their cohabitants register online. A shorter questionnaire was then sent repeatedly to all participants every 4 weeks for an update on their symptoms, COVID-19 exposure, and testing. All individuals at risk of positivity to SARS-CoV-2 infection and their cohabitants have been invited for a nasopharyngeal swab molecular test and a serum antibody test at the CHRIS study center.

*Stage 3:* To trace and monitor antibody response over time, all individuals testing positive to either the nasopharyngeal or the serum test in Stages 1 or 2 have been invited to repeat the serum antibody test every three months for a year, since their first measurement.

As both part of stage 1 and 2, the Limesurvey baseline questionnaire was accessible online between July 2020 and August 2021. An invitation was sent to all participants of the CHRIS baseline study in July 2020 and a reminder in January 2021, with ten access tokens for the participants and their cohabitants. Each person had to register first by inserting access token, personal data, contact information, and filling out the online informed consent form. Afterwards, they were asked to fill-out the baseline questionnaire.

The CHRIS COVID-19 participants who were invited to the blood drawing for serum antibody screening without having filled-out a baseline questionnaire were asked to fill out at the CHRIS study center a paper-based version of the questionnaire.

The baseline questionnaire full text is available at J:\5-5 New Research Initiatives\CHRIS-Covid-19\Questionnaires.

## 2. History version changes

**Data cleaning process:**

**variables added:** ccbq02a ccbq03a ccbq04a ccbq11 ccbq20 ccbq27a ccbq27b

**other:** ccbq02 - free text was categorized.

### **3. Data cleaning**

1. The CHRIS COVID-19 baseline questionnaire dataset was loaded in Stata.
2. Sex (sex) and date of birth (bdate) were generated from the fiscal code as auxiliary variables.
3. The submission date variable ccbqdate was generated from the Limesurvey timestamp. The timestamps were deleted from the dataset. For random-sample participants with paper-based questionnaires, the participation date from the biochemical trait dataset was used instead.
4. The age at participation, fixed for everyone at February 1<sup>st</sup>, 2020, was calculated as (ccbqdate-bdate)/365.25, and its integer part was saved as cc\_age.
5. The language used variable ccbqlang was generated from the Limesurvey startlanguage information and the filename information of the scanned paper-based questionnaires.
6. The respondent variable ccbqresp was generated from the main respondent information.
  - a. For random-sample participants, missing values were set to 1 ("For myself");
  - b. Missing values and 1 were set to "For a minor person or a person with legal guardian living in my household" (3) for age below 14 years old;
  - c. Missing values and 2 were set to "For a minor person or a person with legal guardian living in my household" (3) for age below 18 years old;
  - d. 3 was set to "For an adult person living in my household" (2) for age between 18 and 60 years old;
  - e. Missing values were set to "Missing by design" (-99) for ccbqresp=2 or ccbqresp=3;
  - f. Missing values were set to "Unexpected missing" (-89) otherwise.
7. The sub do-file cr-z-do\_corrections\_baseline.do was run to implement the corrections to questionnaire answers reported by telephone or email by participants to the study call center.
8. The value labels were defined and assigned to the variables as appropriate.
9. For the residential municipality variable ccbq02, free text answers were classified according to the scheme used in CHRIS baseline.
10. For the number of household members variable ccbq03:
  - a. Values between 21 and 99 were set to "Out of range" (-86);
  - b. Missing values were set to "Missing by design" (-99) for age below 14 years and for ccbqresp=2 or ccbqresp=3;
  - c. Missing values were set to "Unexpected missing" (-89) otherwise.
11. For the total number of rooms in the house variable ccbq04, missing values were set to:
  - a. Missing values were set to "Missing by design" (-99) for age below 14 years and for ccbqresp=2 or ccbqresp=3;
  - b. Missing values were set to "Unexpected missing" (-89) otherwise.
12. For the doctor-diagnosis of any listed disease variable ccbq11:
  - a. Missing and negative responses were set to "Prefer not to respond/don't know" (-88) if the response option "Prefer not to respond/don't know" was selected;
  - b. Missing values were set to "Unexpected missing" (-89) otherwise;
  - c. Values -89, -88, or 2 were set to "Yes" (1) if any disease was selected.
13. The disease variables ccbq11a to ccbq11r had their observations transformed as follows:

- a. Missing values were set to “No” (2) if a paper-based questionnaire was used and ccbq11=1;
  - b. Missing values and 2 were set to “N/A” (-99) if ccbq11=2 or ccbq11=-88;
  - c. Missing values were set to “Unexpected missing” (-89) if ccbq11=1 or ccbq11=-89.
- 14. For the pregnancy variable ccbq16, all values were set to “Missing by design” (-99) if sex=“Male” and for age outside 14 and 55 years old.
- 15. For the variables ccbq12 to ccbq19:
  - a. Value 3 was transformed to “Prefer not to respond/don’t know” (-88);
  - b. Missing values were set to “Unexpected missing” (-89) otherwise.
- 16. For the medication intake in the last 4 weeks variable ccbq20:
  - a. Missing and negative responses were set to “Prefer not to respond/don’t know” (-88) if the response option “Prefer not to respond/don’t know” was selected;
  - b. Missing values were set to “Unexpected missing” (-89) otherwise;
  - c. Values -89, -88, or 2 were set to “Yes” (1) if any medication was selected.
- 17. For the single drug variables ccbq20a to ccbq20l:
  - a. Missing values were set to “No” (2) if a paper-based questionnaire was used and ccbq20=1;
  - b. Missing values and 2 were set to “Missing by design” (-99) if ccbq20=2 or ccbq20=-88;
  - c. Missing values were set to “Unexpected missing” (-89) if ccbq20=1 or ccbq20=-89.
- 18. For the primary occupation variable ccbq21:
  - a. All values were set to “Missing by design” (-99) for age below 14 years old;
  - b. Value 8 was transformed to “Prefer not to respond/don’t know” (-88);
  - c. Missing values were set to “Unexpected missing” (-89) otherwise.
- 19. For the sector of activity variable ccbq22:
  - a. Missing values were set to “Missing by design” (-99) if responses to ccbq21 other than 1 or 2;
  - b. All values were set to “Missing by design” (-99) for age below 14 years old and for ccbq21=-88 and if the participant was unemployed or inactive ( $3 \leq \text{ccbq21} \leq 7$ );
  - c. Value 14 was transformed to “Prefer not to respond/don’t know” (-88);
  - d. Missing values were set to “Unexpected missing” (-89) otherwise.
- 20. For the smoking status variable ccbq23:
  - a. All values were set to “Missing by design” (-99) for age below 14 years;
  - b. Missing values were set to “I currently smoke” (1) if valid responses to the current daily number of cigarettes (ccbq23a) were given;
  - c. Missing values were set to “I smoked in the past and have now quit altogether” (3) if valid responses to when they quit smoking (ccbq23b) were given;
  - d. Value 5 was transformed to “Prefer not to respond/don’t know” (-88);
  - e. Missing values were set to “Unexpected missing” (-89) otherwise.
- 21. For the current daily number of cigarettes variable ccbq23a:
  - a. Missing values were set to “Missing by design” (-99) for responses to ccbq23 other than 1;
  - b. All values were set to “Missing by design” (-99) for ccbq23 between 2 and 4;
  - c. Missing values were set to “Unexpected missing” (-89) otherwise.
- 22. For the variable on when they quit smoking ccbq23b:

- a. Missing values were set to “Missing by design” (-99) for responses to ccbq23 other than 3;
  - b. All values were set to “Missing by design” (-99) for ccbq23 equal 1, 2, or 4;
  - c. Missing values were set to “Unexpected missing” (-89) otherwise.
23. Afterwards, the smoking status variable ccbq23 was corrected as follows:
- a. All values were set to “I currently smoke” (1) if valid responses to the current daily number of cigarettes question ccbq23a were given;
  - b. All values were set to “I smoked in the past and have now quit altogether” (3) if valid responses to the question ccbq23b on when they quit smoking were given.
24. For the physical activity variable ccbq24:
- a. Value 3 was transformed to “Prefer not to respond/don’t know” (-88);
  - b. Missing values were set to “Unexpected missing” (-89) otherwise.
25. For the self-reported weight (ccbq25) and height (ccbq26) variables:
- a. The values were exchanged in case a weight over 80 kg and height between 10 and 80 cm (for age below 14 years old) or a weight over 150 kg and height below 100 cm (for all participants) was reported.
  - b. For height below 2, as reporting unit meter was assumed, and the value was multiplied by 100;
  - c. For adult participants with reported height below 100 cm and weight below 150kg, 100 cm were added to the reported height.
  - d. In other dubious cases, the consistency with the information collected in the CHRIS baseline study was checked. The values were then either corrected or set to “Out of range” (-86).
26. Body-Mass-Index (BMI) was computed from the height (ccbq26) and weight (ccbq25) variables and stored in a new variable named ccbq27a.
27. The BMI was categorized into four categories (<18.5, 18.5-24.99, 25-29.99, ≥30) and saved as ccbq27b.
28. The baseline information dataset (ccbq\*) and the variables of the baseline questionnaire which were included in follow-up questionnaire (cclq\*) were saved in different files.

Stata v16.1 was used for the data cleaning process. The do-file is available at the [LINK-TO-cr\\_06b\\_clean\\_BQ\\_limesurvey-paper.do](#).

In cr\_09\_IC\_household.do (do-file not available because it includes personal data of the participants), variables with imputed information for residential municipality (ccbq02a), number of household members (ccbq03a), and total number of rooms in the house (ccbq04a) were computed using the information reported by other household members and added to the dataset.

#### 4. Data structure

The variables listed in table 1 constitute all the variables associated with the survey responses.

**Table 1. Baseline information variables list**

Variable	Description	Unit of reference	Coding	Filter	Notes	Version	Available	Derived
ccbq01	Educational qualification		1 No title 2 Primary school 3 Lower secondary school 4 Vocational school 5 Certificate of upper secondary school 6 Bachelor's degree or equivalent 7 Master or higher degree	age>=14		1,2,3,4	Yes	No
ccbq02	Residential municipality		1 GLURNS 2 GRAUN I. V. 3 KASTELBELL-TSCHARS 4 LAAS 5 LATSCH 6 MALS 7 MARTELL 8 PRAD A. ST. 9 SCHLANDERS 10 SCHLUDERNS 11 SCHNALS 12 STILFS 13 TAUFERS I. M. 21 Burggrafenamt district 22 Other South Tyrol 31 Italy 32 Austria 33 Germany 34 Switzerland / Liechtenstein 42 Spain / Portugal	ccbqresp=1	Derived from categorical response options (1-13) and free text classification of the OTHER category	1,2,3,4	Yes	Yes
ccbq02a	Residential municipality		1 GLURNS 2 GRAUN I. V. 3 KASTELBELL-TSCHARS	ccbqresp=1	Derived from categorical response options (1-13) and free text classification	1,2,3,4	Yes	Yes

Variable	Description	Unit of reference	Coding	Filter	Notes	Version	Available	Derived
			4 LAAS 5 LATSCH 6 MALS 7 MARTELL 8 PRAD A. ST. 9 SCHLANDERS 10 SCHLUDERNS 11 SCHNALS 12 STILFS 13 TAUFRERS I. M. 21 Burggrafenamt district 22 Other South Tyrol 31 Italy 32 Austria 33 Germany 34 Switzerland / Liechtenstein 42 Spain / Portugal		of the OTHER category Missing responses of household members were imputed with the place of residence information provided in the CHRIS COVID-19 registration form or in the CHRIS contact database			
<b>ccbq03</b>	Number of household members, including yourself	persons	<integer>	ccbqresp=1		1,2,3,4	Yes	No
<b>ccbq03a</b>	Number of household members, including yourself	persons	<integer>		Missing responses of household members were imputed with the median of the valid responses of participants from the same household	1,2,3,4	Yes	Yes
<b>ccbq04</b>	Total number of rooms in the house		1 '1' 2 '2' 3 '3' 4 '4' 5 '5' 6 '6 or more'	ccbqresp=1		1,2,3,4	Yes	No
<b>ccbq04a</b>	Total number of rooms in the house		1 '1' 2 '2' 3 '3' 4 '4' 5 '5' 6 '6 or more'		Missing responses of household members were imputed with the median of the valid responses of participants from the same household	1,2,3,4	Yes	Yes
<b>ccbq11</b>	Has a doctor ever diagnosed you with any of the following diseases?		1 Yes 2 No		Derived from exclusive response options 'None of	1,2,3,4	Yes	Yes

Variable	Description	Unit of reference	Coding	Filter	Notes	Version	Available	Derived
					the above' and 'Prefer not to respond/Do not know'			
<b>ccbq11a</b>	Diabetes		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11b</b>	Other metabolic dysfunction		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11c</b>	Respiratory allergy		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11d</b>	Other type of allergy		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11e</b>	Asthma		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11f</b>	Chronic bronchitis		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11g</b>	Other lung disease (excluding cancer)		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11h</b>	Hypertension		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11i</b>	Arrhythmia		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11j</b>	Ischemic or cerebrovascular disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11k</b>	Other cardiovascular disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11l</b>	Kidney disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11m</b>	Liver disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11n</b>	Autoimmune rheumatic disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11o</b>	Musculoskeletal disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11p</b>	Blood disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11q</b>	Mental or affective disease		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No
<b>ccbq11r</b>	Cancer		1 Yes 2 No	ccbq11=1		1,2,3,4	Yes	No



Variable	Description	Unit of reference	Coding	Filter	Notes	Version	Available	Derived
<b>ccbq12</b>	Have you been hospitalized in the past 12 months?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq13</b>	Have you undergone surgery under general anaesthesia in the past 12 months?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq14</b>	Have you undergone chemo or radiation therapy in the past 12 months?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq15</b>	Have you ever had an organ transplant, in your lifetime?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq16</b>	Are you currently pregnant or have you been pregnant over the past 6 months?		1 Yes 2 No	sex=2 & age>=14 & age<=55		1,2,3,4	Yes	No
<b>ccbq17</b>	In the past 12 months, have you received a vaccine for seasonal flu?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq18</b>	In the past 12 months, have you received a vaccine for pneumococcal infection?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq19</b>	In the past 12 months, have you received any other vaccinations?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq20</b>	In the past 4 weeks, have you regularly taken any of the following medications?		1 Yes 2 No		Derived from exclusive response options 'None of the above' and 'Prefer not to respond/Do not know'	1,2,3,4	Yes	Yes
<b>ccbq20a</b>	Use of aspirin		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20b</b>	Use of antithrombotics		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20c</b>	Blood pressure lowering therapy		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20d</b>	Cholesterol lowering therapy		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20e</b>	Use of diabetes medication		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20f</b>	Use of thyroid disease drugs		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20g</b>	Hormone therapy		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No

Variable	Description	Unit of reference	Coding	Filter	Notes	Version	Available	Derived
<b>ccbq20h</b>	Use of Asthma-bronchitis medication		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20i</b>	Use of cortisone or immunosuppressive drugs		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20j</b>	Use of pain-relief drugs (NSAID)		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20k</b>	Use of anxiety, sedative, or anti-depressant drugs		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq20l</b>	Use of vitamin D		1 Yes 2 No	ccbq20=1		1,2,3,4	Yes	No
<b>ccbq21</b>	In the 12 months before the coronavirus-related emergency, what was your primary occupational status?		1 Stable employment 2 Temporary or seasonal employment 3 Unemployed seeking employment 4 Student 5 Retired 6 Housekeeper 7 Other	age>=14		1,2,3,4	Yes	No
<b>ccbq22</b>	What was your main sector of activity?		1 Agriculture, forestry and livestock farming 2 Mining, construction and industry 3 Craftsmanship 4 Provision of essential services 5 Transport and delivery 6 Accommodation and catering 7 Trade 8 Healthcare 9 Social services 10 Education 11 Public administration 12 Police, armed and public security forces 13 Other	ccbq21=1 or ccbq21=2		1,2,3,4	Yes	No
<b>ccbq23</b>	Do you smoke cigarettes (or tobacco in general)?		1 I currently smoke 2 I occasionally smoke 3 I smoked in the past and have now quit altogether 4 I have never smoked	age>=14		1,2,3,4	Yes	No

Variable	Description	Unit of reference	Coding	Filter	Notes	Version	Available	Derived
<b>ccbq23a</b>	How many cigarettes per day?		1 up to 10 2 between 10 and 20 3 over 20	ccbq23=1		1,2,3,4	Yes	No
<b>ccbq23b</b>	How long ago?		1 less than 1 month 2 between 1 and 6 months 3 at least 6 months	ccbq23=3		1,2,3,4	Yes	No
<b>ccbq24</b>	Do you engage in moderate or vigorous physical activity, 2 days or more per week for at least 5 hours in total?		1 Yes 2 No			1,2,3,4	Yes	No
<b>ccbq25</b>	What is your weight?	kg	<float>			1,2,3,4	Yes	No
<b>ccbq26</b>	What is your height?	cm	<integer>			1,2,3,4	Yes	No
<b>ccbq27a</b>	Body-Mass-Index	kg/m <sup>2</sup>	<float>		Derived from ccbq25 and ccbq26	1,2,3,4	Yes	Yes
<b>ccbq27b</b>	Body-Mass-Index category		1 <18.5 2 18.5 - <25 3 25 - <30 4 30+		Categorization of ccbq27a	1,2,3,4	Yes	Yes
<b>ccbqdate</b>	Submission date of CHRIS COVID-19 baseline questionnaire		<date>				Yes	No
<b>ccbqlang</b>	Language used in CHRIS COVID-19 baseline questionnaire		1 German 2 Italian				Yes	No
<b>ccbqresp</b>	For whom are you filling out this questionnaire?		1 For myself 2 For an adult person living in my household 3 For a minor person or a person with legal guardian living in my household				Yes	No
<b>ccbqver</b>	Version of CHRIS COVID-19 baseline questionnaire		1 Paper-based, random sample data collection 2 Limesurvey, CHRIS-Covid-19 Baseline v1.0 3 Limesurvey, CHRIS-Covid-19 Baseline v1.1 4 Limesurvey, CHRIS-Covid-19 Baseline v1.2				Yes	No

## 5. Advices for the analysis

The analyst should keep in mind that the participation to the baseline questionnaire was open during the whole screening study period between July 2020 and August 2021. Questionnaires could be fill-out both for the participant themselves and as a proxy for cohabitants.

For participants to the random sample examination, a paper-based questionnaire was available at the study point in case the baseline questionnaire had not been filled-out online.

Cohabitants of potentially positive individuals (either because of a contact or because of a self-reported positive test) were also invited to the serum antibody screening and could fill -out the baseline questionnaire at the study center if they had not yet participated to the online survey.

## 6. References

- Pattaro C, Barbieri G, Foco L, Weichenberger CX, Biasiotto R, De Grandi A, Fuchsberger C, Egger C, Amon VSC, Hicks AA, Mian M, Mahlknecht A, Lombardo S, Meier H, Weiss H, Rainer R, Dejacó C, Weiss G, Lavezzo E, Crisanti A, Pizzato M, Domingues FS, Mascalzoni D, Gögele M, Melotti R, Pramstaller PP. Prospective epidemiological, molecular, and genetic characterization of a novel coronavirus disease in the Val Venosta/Vinschgau: the CHRIS COVID-19 study protocol. *Pathog Glob Health*. 2022 Mar;116(2):128-136. doi: [10.1080/20477724.2021.1978225](https://doi.org/10.1080/20477724.2021.1978225). PubMed PMID: [34637685](https://pubmed.ncbi.nlm.nih.gov/34637685/)
- Barbieri G, Pizzato M, Gögele M, Giardiello D, Weichenberger CX, Foco L, Bottigliengo D, Bertelli C, Barin L, Lundin R, Pramstaller PP, Pattaro C, Melotti R. Trends and symptoms of SARS-CoV-2 infection: a longitudinal study on an Alpine population representative sample. *BMJ Open* 2023;13:e072650. doi: [10.1136/bmjopen-2023-072650](https://doi.org/10.1136/bmjopen-2023-072650)