Annex 6

Unit non-response and item non-response

(paragraph 6.3.3.1)

General note: All questions are mandatory to be filled in except those explicitly marked as 'Optional question'.

Unit non-response	**Answers**
A) Non-response rate (total and for each mode of data collection; where substitutions are made in cases of unit non-response, non-response rates should be calculated before and after substitutions; including initial selected units and all substitutes for the calculation of 'Non-response rate after substitution').	27 600 individuals were sampled, 14 192 responded and 11 430 belonged to the target population but did not respond : the global response rate is therefore 55,4 % CATI data collection : 19 359 individuals were contacted by phone, 8 091 responded directly by phone and 10 018 did not respond. The response rate for CATI data collection is 44,7 % CAPI direct data collection : 8 241 persons were directly concerned by face-to-face interviews. 4 918 are respondents and 2 765 are nonrespondents. The response rate for direct CAPI data collection is 64,0 % CAPI data collection after CATI : 3 216 sampled persons initially contacted by phone were also contacted for face to face interviews once the cati data collection process was complete for them without the possibility of obtaining any contact with the sampled person. These persons belong to the 10 018 nonrespondents of cati data collection. 1 183 are respondents and 859 nonrespondents, their response rate is therefore 57,9 %

	Each person sampled received before the start of data collection an official letter describing the survey and giving also mandatory legal information to the sampled person. The letter also indicated the references of the survey's website and the address (phone and mail and website) of a hotline sampled person could contact to schedule their interview
B) Methods used for reducing unit non-response (where applicable, give a description of measures taken during the fieldwork to reduce the non-response; e. g. advance notification in form of letters or phone calls, system of reminders, number of visits, number of attempts for phone calls, etc.).	<u>Cati data collection</u> : For each person and each phone number available for her, the minimum number of calls was 25 if the phone number was associated with the person in the sampling frame, or 10 if it was associated with the household or another member of the household, The data collection process had to respect various conditions : calls for a same phone number could not be too close and should be made as various times and days in the weeks. Especially, calls could be made from Monday to Saturday. For each sampled individual and each phone number, at least 5 calls should be passed on Saturday and at least 5 phone calls between Monday and Friday in the evening. If a person was contacted and refused to answer, its phone number was identified and transmitted to a pool of experienced interviewers who tried to contact her again. The different rules set on the cati data collection process were hard to follow for the firm in charge of the data collection process, so we followed precisely the advancement of the data collection process with weekly meetings during which priority rules were decided on the parts of the sample on which to concentrate the data collection effort. Finally, as already mentioned, a random half of the person that could not be reached by phone

	were transmitted to the face to face interviewers. <u>Capi data collection</u> : For each address, at least 5 attempts of contact should be made, split on at least three weeks, at least one the weekend (Saturday), and at least one in the evening. At each attempt of contact, the interviewer had to leave a letter in the mailbox reminding the sampled person about the survey. For all data collections, letters and mails were also sent at call back to remind the nonrespondents that the survey was mandatory. A final letter, reminding them that due the fact the survey is mandatory, they could face the payment of a fine in case of nonresponse, was finally sent.
C) Description of the information available on non-response cases (e. g. demographic and geographic information, health information, etc.).	Information available on nonresponse and usable for nonresponse treatment are variables coming from the fiscal data files used to build the sampling frame : detailed information on incomes (nature of incomes and amount), on the dwelling inhabited at the time of the survey (surface, number of rooms, nature of equipment…), on the other persons living in the same dwelling and / or belonging to the same fiscal household (sex, age, relationship with the sampled person…). Information on the elements of social redistribution the sampled person benefits from are also available. Information are however not all of the same quality. Due to the fact data come from the fiscal administration, their quality depends on the fact they are used to determine the amount of taxes each one has to pay. Sex and ages are for instance not always available, especially for persons aged less

	than 18 years old.
D) Description of the bias risks associated with non-response.	Nonresponse to the survey was rather high, higher than expected, especially on the cati data collection. The risks associated with nonresponse are classic : nonresponse could causes biases in the estimates using the survey data if the respondents differ from nonrespondents on dimensions that nonresponse treatments could not take into account. If for instance persons with severe impairments are harder to reach than persons presenting no impairments with the same level and nature of incomes, they will be underweighted in the sample. A matching of EHIS wave 3 sample with data for the French Système national de données de santé (national system of health data), containing all data on treatments payed for by the French Social Security, has been realised, later than initially planned due to the situation in link with covid-19. We will start exploiting it in November 2021. These data contain information on health consumptions 10 years before the survey data collection, and will be enriched each year during five years with data on health consumptions. This will help us check whether there exists biases in the survey sample on health consumptions and also assess the quality of responses to Ehis questionnaire on health consumptions.
E1) Indication whether substitution used here (only one answer accepted / possible).	□ Yes x No
E2) *Optional question :* If 'Yes' is marked, what method was used (stratified oversampling, other method)	
E3) *Optional question :* If 'Yes' is marked, details of the method used including the criteria for substitution: - which non-responding units were substituted (e.g. non-contacts, refusals, other non-respondents, ineligible units, etc.)	

- what stage the non-responding units were substituted - which criteria were chosen for the selection of substitute units (characteristics of respondents taken into account).	
E4) *Optional question :* If 'Yes' is marked, indication of the substitution rate (ratio of substituted units to total sample)	
Item non-response	**Answers**
A) Item non-response rates (average – minimum – maximum) across the health variables. **Note**: * The item non-response rates are to be calculated un-weighted and before imputation; items for which a proxy respondent gave an answer but is not recommended according to the EHIS manual should be excluded from the calculation (i.e. from numerator and denominator). * If proxy interviews were allowed in the national survey beyond the recommendations in the EHIS manual, item non-response could be calculated, in addition to the calculation method recommended above, also including proxy interviews.	Nonresponse rates are computed as the number of nonresponse for a question divided by the number of persons that should have answered the questions. If proxies are regarded are nonrespondents: Max : 32,9 % Mean : 4,1 % Min : 2 % If proxies are regarded as respondents : Max : 2,5 % Mean : 0,3 % Min : 0 %
B) Total item non-response (un-weighted; number of completed values divided by expected values over all variables of all respondents; taking into account filtering and derogations).	If responses given by proxies are not taken into account : 2,8 % If responses given by proxies are taken into account : 0,3 %
C) Variables for which the item response (i.e. before imputation) was below 90%. Indication of the codes of variables, including common variables.	No variable concerned
D) *Optional question :* References to methodological notes and results of non-response analysis or other methods to assess the effects of non-response in the national survey.	