EHIS_ESQRS_A_EL_2014_0000

National Reference Metadata in ESS Standard for Quality Reports Structure

(ESQRS)

Compiling agency: HELLENIC STATISTICAL AUTHORITY ELSTAT
Time Dimension: 2014-A0
Data Provider: EL1
Data Flow: EHIS_ESQRS_A:1.0

1. Contact	Top

1.1. Contact organisation : HELLENIC STATISTICAL AUTHORITY ELSTAT

1.2. Contact organisation unit : GENERAL DIRECTORATE OF STATISTICAL SURVEYS POPULATION AND LABOUR MARKET STATISTICS DIVISION SPECIAL HOUSEHOLD SURVEYS SECTION (G53)

1.3. Contact : 18510, PIREOS 46 & EPONITON STR, PIRAEUS , GREECE

1.4. e-mail address: m.chalkiadaki@statistics.gr

2. Introduction	Top

The production of the quality reports is part of the implementation of the European Health Interview Survey (EHIS) instrument. In order to assess the quality of data at national level and to make a comparison among countries, the National Statistics Institutes give detailed information mainly on: the entire statistical process, sampling and non-sampling errors and potential deviations from standard definition and concepts.

This document follows the ESS standard for quality reports structure (ESQRS), which is the

main report structure for reference metadata related to data quality in the European Statistical System. It is a metadata template, based on 13 main concepts, which can be used across several statistical domains with the purpose of a better harmonization of the quality reporting requirements in the ESS.

Generally speaking, the EHISis part of a broader community statistical program in which all EU Member States participate and aims to study and provide valuable information as regards the population health status and the determinants affecting it, both at European and National level.

More specifically data will be collected as regards:

  The demographic characteristics (gender and age), educational level and employment

• Health status (general health status, health problems, accidents and injuries, absence from work due to health problems, physical and sensory functional limitations, personal care activities, household activities, pain, mental health)

• Use of Health Care Services (inpatient and outpatient hospital care, use of ambulatory and home care), visits/consultations to family doctors/general practitioners/medical or surgical specialists/ dentists, use of medicines, preventive services (flu vaccination, measuring blood pressure and blood cholesterol level, etc.), unmet need for health care

• Factors that influence health status –positively or negatively– such as physical exercise, healthy food consumption, smoking consumption of alcoholic beverages etc.

• The net monthly income of the household

The legal framework for developing the EHIS is the Regulation No **1338/2008** of the European Parliament and of the Council of 16 December 2008 on public health and safety at work. In addition, detailed specification of the data and metadata are provided pursuant to the Commission Regulation **No 141/2013 of 19 February 2013**.

3. Quality management - assessment	Top

Quality assurance	Quality assessment
	Quality assessment is carried out by ELSTAT and Eurostat. The quality of the survey is ensured by the existence of the methodological handbook issued by Eurostat, as well as by the use of common variables, in order to improve comparability of results in all member states Quality is also ensured by the implementation of the European Statistics Code of Practice. The sample size is such as to ensure high accuracy results, representative for the reference population and all necessary steps are taken so as to make all appropriate checks and minimize measurement errors in data collection.

4. Relevance	Top

4.1. Relevance - User Needs

Main users of the survey data is Eurostat, OECD, Universities (professors, graduate and post graduate students), researchers on the field of health etc. In general the users' requests are being satisfied.

4.2. Relevance - User Satisfaction

The Statistical Information and Dissemination Section of ELSTAT conducts a survey on users satisfaction: http://www.statistics.gr/user-satisfaction-survey

In order to acquire extra health data for national needs, a Working Group was established in ELSTAT for the preparation of the questionnaire. Experts from the Ministry of Health,

Universities and the National Confederation of Disabled participated in and a sufficient number of questions have been included in the questionnaire.

All extra questions and the modifications of existing questions are included in Annex 3: National adaptations of the model questionnaire.

In order to promote the data use, the produced data sets disseminated as follows:

- Tables and metadata (statistical and methodological contents, questionnaires, etc) posted on the Internet: www.statistics.gr (ELSTAT website)
- Press Releases
- a users' micro data file being constructed by using the anonymisation criteria.

4.3. Completeness

The completeness and breakdowns of data are considered to be very satisfactory according to the requirements set out in Commission Regulations 141/2013 and 1338/2008.

4.3.1. Data completeness - rate

The completeness of data produced is considered very satisfactory according to the above Regulations and the national experts participated in the WG.

5. Accuracy and reliability

The concept of accuracy refers to the precision of estimates computed from a sample rather than from the entire population. Accuracy depends on sample size, sampling design effect and structure of the population under study. In addition to that, sampling errors and non sampling errors need to be taken into account. Sampling errors refer to the variability that occurs at random way because of the use of a sample rather than a census. Finally, the non – sampling errors are errors that occur in all phases of the data collection and production process.

5.1. Accuracy - overall

The sample size has been defined from Organization and Methodology Division according to the provisions set out in the survey implementing Regulation, allowing the production of high accuracy estimations for the Country as a whole and for the Great Geographical Areas of Greece (NUTS-1 level).

Sampling errors have been calculated from the competent Division and presented below (5.2).

The non-sampling errors are divided into the following categories:
1. Coverage error
2. Measurement error
3. Processing error
4. Non-response error

5.2. Sampling error

Please describe the calculation of the design effect and standard error. If outliers received special treatment in estimation (e.g. if they were excluded from estimation), it should be clearly described.	As a multistage stratified sampling design was applied, the variance estimation procedure pools stratum variance estimates to compute the overall variance estimate. In each stratum, the variance calculates the variability of all estimates coming from each separate primary sampling unit. For a multistage sample design, in which the first-stage sampling fraction is small and the first-stage sample

is drawn with replacement and probabilities proportional to their sizes the variance estimation method depends only on the first stage of the sample design. As a result, the required input for variance estimates includes only the first-stage cluster (PSU) and first-stage stratum identification. The design effect was estimated as the ratio of the variance of the actual sampling design (multistage stratified sampling design) to the variance from a hypothetical SRS without replacement of the same size.

Indicator / sub-indicator (variable(s) from which the indicator is derived)	Number of respondents - *n* (unweighted)	Estimated proportion - *p* (weighted)	Standard error - *SE* (with respect of sampling plan)	95% confidence interval | lower; upper |	Design effect deff (if applicable/available)
Respondents aged 15 years or over in good or very good health (HS1)					
All	**5419**	**74.8**	0.6	73.6 - 75.9	1.467
Women	**3020**	**71.0**	0.8	69.4 - 72.6	1.354
Men	**2399**	**78.9**	0.9	77.2 - 80.6	1.636
Respondents aged 15 years or over with a longstanding illness or health problem (HS2)					
All	**4589**	**49.7**	0.8	48.1 - 51.3	2.137
Women	**2897**	**54.1**	1.0	52.2 - 56.1	1.754
Men	**1692**	**44.8**	1.1	42.6 - 47.0	1.943
Respondents aged 15 years or over that were severely limited in activities people usually do because of health problems for at least the past 6 months (HS3)					
All	**1063**	**10.3**	0.4	9.5 - 11.1	1.532
Women	**681**	**11.4**	0.5	10.3 -12.5	1.257
Men	**382**	**9.1**	0.6	8.0 - 10.3	1.632

Respondents aged 15 years or over declaring having been hospitalized in the past 12 months (HO1) (men and women)	**904**	**9.7**	0.4	8.9 - 10.4	1.384
Respondents aged 18 years or over who are obese (BMI>=30, where BMI = BM2 in kg / (BM1 in m * BM1 in m) (men and women)	**1361**	**17.4**	0.6	16.3 - 18.5	1.662
Comments related to the table	The indicators refer to population 15 years or over if not stated else.				

5.2.1. Sampling error - indicators

[not requested]

5.3. Non-sampling error

The non-sampling errors are errors that occur in all phases of the data collection and are divided into the following categories:

1. Coverage error

The specific kind of error due to divergences between the target population and the sampling frame.

2. Measurement error:

They occur at the time of data collection from the questionnaire, the interviewers, and the respondents which due to the sensitive subjects the survey is dealing with, do not provide certain information such as mental health problems or other diseases.

In order to reduce these errors the following actions were done:

- For building up the questionnaire we adopted the questionnaire proposed from Eurostat as the basis. The structure of the questionnaires is similar. The majority of the questions are almost literally copied and translated. In order to finalize the questionnaires, we took into account any observations made on the questionnaires of the surveys' previous wave (2009) together with all comments made by health experts in the working group established in ELSTAT for the survey.
- Detailed guidelines handbook was provided to all interviewers. Also, a second manual on the use of the data entry programs / electronic questionnaires designed using Oracle - SQL (in portable PCs or in central office desktops).
- Training both to interviewers as well as to ELSTAT's personnel-supervisors in charge of the survey in all the regional statistical offices. Training took place mainly in Athens, but whenever this wasn't feasible by Skype). In the training also participated health experts from the Ministry of Health.
- Checks (validation/ logical/ completeness/ flow, etc.) were carried out both by ELSTAT and EUROSTAT (validation checks). Whenever necessary, households were called back from the supervisors. Supervisors were in close communication with interviewers for attending work progress or to discuss and solve occurring problems.

- Participation in the survey of skilled and experienced, with Wave I and other household surveys, personnel and interviewers.

3. Processing error

Quality and quantity checks were made on the data base for correcting any processing error.

Processing error: Occur in post-data-collection processes such as data entry, keying, editing and weighting

- Data entry checks were made from ELSTAT's experienced personnel and also using ESTAT's validation rules
- Longitudinal checks of data from previous Wave
- Codification of questions relating to occupation (ISCO-08), economic activity of the local unit (NACE rev.2).

4. Non-response error (see 5.3.3)

5.3.1. Coverage error

A) Overall quality of the sampling frame. Shortcomings in terms of timeliness (e.g. time lag between last update of the sampling frame and the moment of the actual sampling), geographical coverage, coverage of different subpopulations (institutionalised persons), multiple listings, etc.	**Coverage error : it** is caused by the imperfection of the sampling frame of the 2011 Census of Population being somehow outdated thus resulting in housing units found to be empty or to be used for other purposes, secondary residences, etc. • Ineligible sample cases / out-of-scope units = **62** • households (house/building not existing, no one living in the building/on the address=**21** and • change of residence for a household, selected individual dead between the reference data of the sampling frame and the moment of the interview, etc. =**41**.
B) Actions taken for reduction of the shortcomings of the sampling frame (e.g. corrections for sampling frame under-coverage, dealing with overcoverage and multiple listings, etc.).	n.a

5.3.1.1. Over-coverage - rate

[not requested]

5.3.2. Measurement error

Proxy interviews	Select from here	Answers
A1) Please indicate whether the proxy interviews were used	NO YES	**Yes**
A2) If Yes, for what reasons the proxy interviews were allowed?	Respondents: 1. Suffering from long-term cognitive impairment; 2. Suffering from long-term severe debilitation; 3. Suffering from a long-term sensory impairment that	**1.** Suffering from long term cognitive impairment **2.** Suffering from long term severe debilitation **3.** Away from the household for educational or work

	prevents the interaction between interviewer and interviewee; 4. In hospital/health or social care facility for the entire period of the fieldwork; 5. Away from the household for educational or work purposes for the entire period of the fieldwork in their area of residence. 6. Other (please specify):	purposes for the entire period of the fieldwork. **4.** Hospitalized for health problems for the entire period of the fieldwork.
A3) If Yes, for which part of the questionnaire the proxy interview was allowed?	1. Whole questionnaire 2. Limited to the questions specified in Eurostat guidelines 3. Limited to other questions (Specify which EHIS ones: …) 4. Other, *please specify:*	The whole questionnaire was allowed to be completed in proxy interviews. In the micro data files sent to ESTAT, for 280 cases, answers for the 57 variables at which proxy was not allowed have taken value -1 (together with other missing).
A4) If Yes, please provide the percentage of proxy interviews (in total and for each reason for proxy): (Note: An interview is considered as a proxy interview even if proxy respondents were allowed to answer only a selected number of questions.)	….. (in %)	• 3.4% of sample cases (unweighted) • 5.5% (weighted)

5.3.3. Non response error	

Unit non-response	Select from here	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 will be calculated before and after substitutions.) (For the calculation of 'Non-response rate after substitution': Do not forget to include initial selected units and all substitutes.)		Non response rate = 16.7% after substitutions and Response rate 83.3%= 8223/(9479+395) Non response rate = 17.1% before substitutions and Response rate 82.9%= 8185/(9479+395) *8223-38=8185
B) Methods used for reducing unit non-response Where applicable, give a description of measures	Examples: advance notification in the form of a letter or phone call, system of reminders, number of visits, number of attempts for phone calls, etc.	• An informative letter was sent, almost one month before the survey conduct, in order to inform the households/individuals that

taken during the fieldwork to reduce the non-response		they have been selected for the survey, furthermore asking for their cooperation. • Whenever the interviewer couldn't find the interviewee during his/her visit, he/she left a note with his/her telephone number and the date of the next visit (at least three visits for each surveyed household were made). • Substituting the household members that couldn't be found is –in general- not recommended/allowed. However, such provision existed in the questionnaire only for cases of temporary absence (and after at least of three attempts) or denial of the interviewee for which no other present member was capable to provide information (proxy interview). Totally 38 substitutions (0.46% of effective interviews) have been made, with the use of random numbers. A table of random numbers was given to the interviewers in order to select among the household members the one to provide information at individual level. The same table has been used for the substitution. More specifically, the initially selected individual was conceivably exempted, thus resulting in a decreased number of total household members, including all others but that one, and restarting the selection. The respective cases can be identified in our data files.
C) What information on	Examples: demographic and	Demographic information is

non-response cases is available?	geographic information, health information, etc.	not available. Geographic information for 1713 households (units) not responded and the initially drawn sample is found in the attached table " Geographic information_unit non response".
D) Bias risks associated with non-response		The effect of the non-response on the produced statistics is that it increases variability and introduces bias. Bias is introduced by the fact that the non-respondents may have different survey characteristics (and consequently different values) than the respondents. Generally, the total bias due to non-response is approximately equal to zero if either the response rates or the means of respondents do not vary between strata.
E1) Please indicate whether substitution was used	NO YES	Yes, substitution of individuals.
E2) If Yes, which method was used?	- Stratified oversampling - Other	Stratified oversampling
E3) If Yes, please describe the method used including the criteria for substitution: - which non-responding units are substituted (e.g. non-contacts, refusals, other non-respondents, ineligible units, etc.); - at what stage they are substituted; - the criteria for the selection of substitute units (characteristics of respondents taken into account).		Substitution of individuals for cases of temporary absence (and after at least of three attempts) or denial of the interviewee for which no other present member was capable to provide information (proxy interview). Totally 38 substitutions (0.46% of effective interviews) have been made. The substitution took place immediately after the initial selection of the member with the use of random numbers.
E4) Substitution rate (Ratio of substituted units to total sample)		0.46%

Item non-response	Select from here	Answers
A) Item non-response rate across the health variables (un-weighted and	Average: Min.:	Item non-response = 0%

before imputation) Items for which a proxy gave an answer but is not recommended according to the EHIS manual should be excluded from calculation of item non-response (i.e. from the numerator and denominator). Note: If proxy interviews were allowed beyond the recommendations as stated in EHIS manual, item non-response could be calculated, in addition to the calculation method recommended above, also including proxy interviews.	Max.:	
B) Total item non-response (un-weighted) (number of completed values / expected values over all variables of all respondents taking into account filtering and derogations granted) (link to data completeness rate / substitution rate)		Total item non-response = 0%
C) For which variables the item response (i.e. before imputation) was below 90 %? (Codes of variables, including core social variables)		None
D) References to methodological notes and results of non-response analysis or other methods to assess the effects of non-response.		Re-weighting was applied to amend suitably the extrapolation factors, by taking into account the response rates in all final strata. In this way there is compensation for non-responses, and reduction of bias in the estimation of the survey characteristics. The aim is to remove non-response bias, but in practice, this is unlikely to be fully achieved.

Annexes:
[Geographic information for households not responded](#)

5.3.3.1. Unit non-response - rate
[not requested]

5.3.3.2. Item non-response - rate
[not requested]

5.3.4. Processing error

Data processing	Select from here	Answers

A) Please describe the data entry and coding control process and tools used (all measures to improve the process, e.g. performing double data entry are to be mentioned). (If applicable, describe for different modes of data collection)		The PAPI – method was used to be interviewed the respondents. The electronic questionnaires were designed using Oracle – SQL program. Data entry and coding control process took place either at the ELSTAT premises in Athens or at the Regional Offices by the interviewers. If this was not possible for a reason the process has been done by the Informatics Division. Data entry procedure was completed at the end of March 2015. Quality and quantity checks were made in the database for correcting any processing errors. For the correction of errors due to data entry and coding the following checks were made: • Completeness checks • Flow checks • Logical checks and finally • Longitudinally checks (between Health surveys data of 2009 & 2014).
		The main detected errors concerned the logical correlation of the replies. For example: • Correlation of demographic characteristics among household members (e.g. age of parents and children) • Correlation and Corrections among first names and sex (typing error) Also incompatibilities were detected in: • The Health Status Module (among the HS1….HS3 questions) • The Health Status Module with the rest of the questionnaire (e.g. list of chronic conditions /personal care activities/household activities/pain/health care /visits to doctors /medicine use etc). • The variables HHACT/HHINACT for the respondent (in the EEY.1 questionnaire) and the variable MAINSTAT (in the EEY.2 questionnaire) – in two separate questionnaires. • The variables of walking in comparison with the answers in the list of chronic conditions _ in particular with the CD1G/CD1H/CD1I questions. • The variables of use of health care (HO1 & HO3: the respondent had admission in hospitals as inpatient or day patient) with variables of unmet needs (UN1A & UN1B = 3

		No need for health care)
		• The variables concerning visits/consultations to dentists, general practitioners or family doctors or specialists (AM2 & AM4) with the variable concerning the use of medicines prescribed by a doctor (MD1)
		• The variable PE1 (Work - related physical activity: mostly sitting or standing) in relation to the respondent's age and health problems.
		The data cleaning was made using specific applications and with the support and application of validation rules recommended by EUROSTAT as well. In addition phone calls and communication were made (either with the interviewer or directly with the household) in cases of extreme values/errors in order to be verified the information provided.
D1) Were there questions asked as open and post coded afterwards? (Concerning for example ISCED, NACE, ISCO)	No Yes	Yes.
D2) If Yes, please list codes of the respective EHIS variables:		The questions concerning the occupation and economic activity were open questions, post coding afterwards, using classifications ISCO 08 for occupation and NACE Rev. 2 for the Economic activity.
D3) If Yes, describe the process of post-coding (was it done by the interviewer, in the regional, central office, what were the methods used, etc.).		Occupation and economic activity were post-coded by experienced ELSTAT personnel.

5.3.4.1. Imputation - rate

[not requested]

5.3.4.2. Common units - proportion

[not requested]

5.3.5. Model assumption error

[not requested]

5.3.6. Data revision

[not requested]

5.3.6.1. Data revision - policy

The revision policy concerns either the survey data or the survey itself (ie. the questionnaire,the sample etc),and takes place having into consideration the users' needs for any further statistical information.

5.3.6.2. Data revision - practice

Data transmitted to Eurostat undergo in detailed verifications by implementing automated

validation procedures at the level of variables and breakdowns. If data inconsistencies are found the reporting countries are asked to verify and revise their results.

5.3.6.3. Data revision - average size
[not requested]

5.3.7. Seasonal adjustment
[not requested]

Dates when each of the phases of the projects started/ended	Start date (month/year)	End date (month/year)
A) Preparation of survey (from designing the questionnaire to fieldwork)	Beginning of year 2014	September 2014
B) Data collection (fieldwork)	October 2014	March 2015
C) Data processing (data entry, validating, editing, imputing, etc.)	March 2015	September 2015
D) Data delivery to Eurostat (dates of first and final data delivery approved by Eurostat)	First delivery of microdata 29/09/2015	Final delivery 5/7/2016
E) Dissemination of national results	The national results are available for dissemination right after the 1rst Press Release of the survey results (9/12/2015).	-

6.1. Timeliness

The time lag between data reference period and data publication is approximately twelve months. The microdata file has been firstly delivered in 29/9/2015 and the quality report available by mid January 2016 (extended time limit).

The first delivery for the microdata was in 29/9/2015. Because of the revision the extrapollating factors, by revising the number of sampled persons aged 65+ years old that belong to the age classes 65-74 and 75+ in the calibration procedure a **second delivery** of the microdata file took place in **14/10/2015**. The final delivery took place in **5 /7/2016** after applying trimming procedure in the extrapolation factors.

6.1.1. Time lag - first result
[not requested]

6.1.2. Time lag - final result
[not requested]

6.2. Punctuality

The survey data have been delivered within the predetermined date.

6.2.1. Punctuality - delivery and publication
[not requested]

Information on the survey (in Greek and in English) has been posted on the website of ELSTAT http://www.statistics.gr/statistics/-/publication/SHE22/2014

Users can find the questionnaires of the survey and surveys' guidelines in national language, the SDMX and short user oriented metadata manuals in Greek and in English.

http://www.statistics.gr/documents/20181/c62bcd62-7769-425a-8dc6-44e1b40d389c (questionnaire EEY1- in Greek)

http://www.statistics.gr/documents/20181/426219aa-3936-4ba0-8391-884262fecbe0 (questionnaire EEY2 – in Greek)

 http://www.statistics.gr/statistics/-/publication/SHE22/2014   (questionnaires HS.1/HS.2- in English)

http://www.statistics.gr/documents/20181/180dec65-aaa7-4ae1-80a0-626d9a01fca4 (guidelines in Greek)

http://www.statistics.gr/documents/20181/fc500869-6342-4d88-809e-9d6d439061d0 (SDMX in Greek)

http://www.statistics.gr/documents/20181/f0236852-da5c-4f12-923b-56836b5ccb28 (SDMX in English)

Publication "The Living Conditions in Greece" provides, both in Greek and English languages, the latest statistics illustrating living conditions in Greece, among which health data –for the time being for the first Wave of the National Health Survey of the year 2009. The publication is updated with the latest data the first Friday of January, March, May, July, September and November, so data on Wave 2 of the National Health Survey of the year 2014 .

Relative link:

http://www.statistics.gr/documents/20181/1216578/LivingConditionsInGreece_0115.pdf/3423a9dc-b9ad-41b9-a2f7-6c507bdf040e

On the 9th December 2015 the first Press Release on 2014 survey's data will be announced and posted on the above website. Soon after that date tables with basic results will be posted. A second Press Release is scheduled for 29th February 2016.

Announcement Calendar for 2015 and 2016 in following links(in Greek & in English):

http://www.statistics.gr/documents/20181/300673/calendar_gr_2015.pdf/a3d5c0e8-637d-4261-9c8a-cc9c695b34ac

http://www.statistics.gr/documents/20181/301069/release_calendar_en.pdf/cde9bc82-d15f-4bb0-9b20-49534002567b

 http://www.statistics.gr/documents/20181/300673/calendar_gr_2016.pdf/68ce517f-aa6b-44a3-b0fb-515db3febaf7

 http://www.statistics.gr/documents/20181/301069/calendar_en_2016.pdf/ba60002a-c437-44d3-9d27-12cb41dc8511

 Microdata are available to users upon request in the competent Unit of Statistical Information using the following

data.dissem@statistics.gr

http://dlib.statistics.gr/portal/page/portal/ESYE/

7.1. Dissemination format - News release

Information on the Health Survey is available on the webpage of the Hellenic Statistical Authority

 http://www.statistics.gr/statistics/-/publication/SHE22/2014

7.2. Dissemination format - Publications

http://www.statistics.gr/documents/20181/1216578/LivingConditionsInGreece_0115.pdf/3423a9dc-b9ad-41b9-a2f7-6c507bdf040e

7.3. Dissemination format - online database

For the time being, on line data base for data access does not exist.

7.3.1. Data tables - consultations

7.4. Dissemination format - microdata access

The micro data files are provided to the users after submitting a statistical data request on line in ELSTAT's website
www.statistics.gr/en/statistical-data-request

7.5. Documentation on methodology

- The present ESQRS will be uploaded on ELSTAT's official website
- A short methodological report orientated to the users will be uploaded on ELSTAT's website and
- The metadata report in Euro - SDMX metadata structure will be uploaded too.

7.5.1. Metadata completeness - rate

7.5.2. Metadata - consultations

7.6. Quality management - documentation

Quality isssues concerning the survey are included in the Quality Report.

7.7. Dissemination format - other

The Press Release for the Health Survey of the year 2014 was announced on the 9th of December of 2015 in ELSTAT's website.
http://www.statistics.gr/statistics/-/publication/SHE22/2014

---

8. Comparability Top

According to the Regulations (EC) No 141/2013 and No 1338/2008of the European Parliament and of the Council on public health and health and safety at work, as regards statistics based on the European Health Interview Survey (EHIS) "the availability of comparable Union-wide data is likely to be of great benefit for health and social policy decisions and for scientific purposes, with the use of common tools allowing data consistency across countries, …".

Comparability of data between Member States is assured by keeping the comparability of methods used and of definitions of variables. This target is much aided by the use of EU proposed questionnaire and of the conceptual guidelines and instructions. Definitions used in the Health Survey of Greece are comparable with Eurostat definitions, as much as possible, after the adaptation of the questions to the national situation.

Comparability over time –among the variables of Wave 1 and Wave 2– is feasible for most variables (as also mentioned in Eurostat's manual / conceptual guidelines).

Comparability among the minimum health module (MEHM), also included in the EU-SILC, as far as definitions used is not an issue, as the questions are identical, the same definitions have been used and the same guidelines have been provided to interviewers. However, in the cohesion section, tables included show differences among the results acquired from the two surveys.

As far as the **reference population** is concerned all citizens officially living at Greek territory (population de facto). The sampling frame of the NHS is derived from the Census Population. This Census includes all private households and their current members residing in the territory independently of any socio-economic characteristics they may have. Persons living in collective households and in institutions are excluded from the target population as well as households having as members diplomatic missioners.

DEFINITIONS used in the Health Survey:

HOUSEHOLD: The definition of **household** that Eurostat recommends is used. Household is defined as a person living alone or a group of people who live together in the same dwelling and share expenditures including the joint provision of the essentials of living. Household members contribute to the expenses of the dwelling or take advantage of it as not having income.

HOUSEHOLD MEMBERS: The criterion for recording household members is the usual residence and the time period of 12 consecutive months.

Individuals usually residing in the dwelling, according to the definition below constitute household and are recorded as household members.

An individual is considered to **usual reside** in a dwelling (household) if he/she uses the dwelling for his/her daily rest/relaxation, independently of temporary absences for recreation, holidays, visits to friends or relatives, work, medical treatment of religious obligations

Or

If the above criteria are not valid the individual is considered to usual reside in his/her legal or registered in authorities dwelling.

Only the following persons are considered as usually residing in the interviewed household:

(1)   Persons having lived in the household for the time period of 12 consecutive months before the survey conduct and

(2)   Persons having moved in the household some time during the last 12 months before the survey conduct, intending to stay for at least a year, continuously.

Hence, soldiers, sailors or persons working far from their dwelling/household for quite a long time or persons hospitalized for health problems the above criteria (1) and (2) for usual residence cannot apply and we consider them to reside in their legal registered in authorities dwelling, that is, the specific dwelling/household. The interview is in these cases done via representative (proxy interviews). For the completeness and accuracy of information collected telephone contact with interviewee is sought for recording or cross checking answers in specific questions.

For students away from home, in another city criteria (1) and (2) will be examined and case by case they may or may not be recorded as household members. If registered, proxy interview will be applied and telephone call for cross checking of answers will follow.

8.1. Comparability - geographical

Because of having been applied common definitions of variables and common methods of data production, between our country and all EU countries,therefore,no geographical comparability over time exists.

8.1.1. Asymmetry for mirror flow statistics - coefficient

[not requested]

8.2. Comparability - over time

The Health survey of the year 2014 is carried out for second time in our country- the 2009 survey was the first one. The data, both for survey years 2009 & 2014, are comparable because of implementing for most variables-common definitions and common methods of data production. For the above mentioned reasons, no problem on comparability over time exists.

8.2.1. Length of comparable time series

[not requested]

8.3. Comparability - domain

[not requested]

9. Coherence

The coherence of two or more statistical outputs refers to the degree to which the statistical processes, by which were generated, used the same concepts and harmonized methods.
A comparison between the Health Survey and EU- SILC MEHM outcomes is presented below in 9.1.

9.1. Coherence - cross domain

General health status indicators are compared with the same indicators calculated from the EU-SILC 2014. The results / differences are presented below. Differences between the concepts / definitions used in the two surveys do not exist hence discrepancies may be attributed to the fact that the two surveys serve different purposes; Health Survey targets health whereas EU-SILC targets household income.

Table 1. Population by GENDER %

	Health Survey 2014	EU-SILC 2014
MEN	47.7	48.4
WOMEN	52.3	51.6

Table 2 Population by AGE GROUP %

	Health Survey 2014	EU-SILC 2014
16-24 for SILC 15-24 for Health Survey	11.6	10.4
25-34	14.6	15.1
35-44	17.7	18.3
45-54	16.9	17.2
55-64	14.6	15.0
65+	24.5	24.1

Table 3 Population by ACTIVITY STATUS %

	Health Survey 2014	EU-SILC 2014
Carries out a job or profession, including unpaid work for a family business or holding, an apprenticeship or paid traineeship, etc.	36.2	37.1
Unemployed	14.7	15.7
Pupil, student, further training, unpaid work experience	8.7	6.5
In retirement or early retirement or has given up business	25.8	25.2
Permanently disabled	2.6	1.2
In compulsory military or community service	0.047	0.1
Fulfilling domestic tasks	11.9	12.9
Other inactive person	0.1	1.1

As far as target populations are concerned, in the Health Survey target population is population aged 15+, while in the EU-SILC population aged 16+.

Table 1.1. General Health (HS1) %

	Health Survey 2014	Cumulative Percent	**EU-SILC 2014**	Cumulative Percent
Very good	**38.8**	38.8	**44.8**	44.8
Good	**36.0**	74.8	**28.8**	73.6
Fair	**18.2**	93.0	**15.7**	89.3
Bad	**5.2**	98.2	**7.8**	97.1
Very bad	**1.8**	100.0	**2.9**	100.0

Higher differences are recorded for variable HS2 on the existence of longstanding illness or health problem. As pre-mentioned this could be attributed to the different purposes the two surveys serve. When you are in the process of answering questions concerning your health – having been informed on the scope and concept of the survey- definitely your health status / health problems (if any) are on your mind, warding off spontaneous answer, which is the case in the EU-SILC MEHM, at which respondents have been prepared to answer questions on their income. The purpose of the variable is to measure the presence of **long-standing** limitations, and although questions HS2 and CD1 are independent check if the answers were coded correctly was encouraged, for such cases appearing in the data. As much as possible, clarification –via telephone calls– was asked from respondents who provided inconsistent answers.

Also it should be noted that in the Greek questionnaire, for national needs, more chronic diseases / health problems have been added in the CD table. Some interviewers probably used this check during the interview and after the CD items asked respondent for clarification.

Table 2.1. Suffering from any longstanding illness or health problem %

	Heath Survey 2014	**EU-SILC 2014**	Difference
Yes	**49.7**	**23.9**	-25.8
No	**50.3**	**76.1**	+25.8

Table 3. Limitations (for at least the past 6 months) in activities because of a health problem. %

	Health Survey 2014	**EU-SILC 2014**	Difference
Severely limited	**10.3**	**11.3**	-1.0
Limited, but not severely	**19.4**	**12.6**	-6.8
Not limited at all	**70.3**	**76.0**	-5.7

Concerning unmet need for health services (experiencing delay or not getting health care at all), comparability is not directly feasible with data from EU-SILC MEHM or with Wave 1 data of the Health Survey (2009). In the EU-SILC MEHM variable records unmet need for health care and not delays (even serious) in getting health care. Also EU-SILC question asks for unmet need for medical care from doctor or surgeon while in the Health Survey unmet need for any health care is recorded. In any case, percentages for unmet need for health from the EU-SILC 2014 and the Health Survey 2014 are 9.7% and 13.6%, respectively.

Concerning unmet need for dental services due to financial reasons, in EU-SILC the percentage of population reporting unmet need due to financial reasons as main reason is 12.4%, while from the Health Survey the respective percentage is 15.2%.

9.1.1. Coherence - sub annual and annual statistics
[not requested]

9.1.2. Coherence - National Accounts
[not requested]

9.2. Coherence - internal
[not requested]

10. Cost and Burden

The Health Survey has been designed to keep respondent burden under control in order to avoid high non-response rate and to ensure good quality of the collected information. Burden concerns mostly the required time, which must be spent by the respondent for data provision. Nevertheless, there is not possibility for any other time relief.
 Cost:

External collaborators	
Type of work	**Total cost**
Sampling frames' updating	27448,54
Data collection	142595,82
Interpeters for data collection	412,8
TOTAL External collaborators	**170457,16**
Staff	
Type of work	**Total cost**
Sampling frames' updating	1042,00
Training seminars	11800,00
Data collection	0,00
TOTAL Staff	**12842,00**

TOTAL COST	**183.299,16**

11. Confidentiality

11.1. Confidentiality - policy

The issues concerning the observance of statistical confidentiality by the Hellenic StatisticalAuthority (ELSTAT) are arranged by articles 6, 7 and 8 of the Law 3832/2010, as amended by article 90 paragraph 8 of Law 3842/2010 and by article 10 of Law 3899/2010, as well as by article 8 of Law 2392/1996, which was brought back into force, in accordance with article 90 paragraph 8 of Law 3842/2010.

Furthermore, ELSTAT disseminates the statistics in compliance with the statistical principles of the European Statistics Code of Practice and in particular with the principle of statistical confidentiality.

11.2. Confidentiality - data treatment

Confidentiality – data treatment

ELSTAT protects and does not disseminate data it has obtained or it has access to, which enable the direct or indirect identification of the statistical units that have provided them by the disclosure of individual information directly received for statistical purposes or indirectly supplied from administrative or other sources. ELSTAT takes all appropriate preventive measures so as to render impossible the identification of individual statistical units by technical or other means that might reasonably be used by a third party. Statistical data that could potentially enable the identification of the statistical unit are disseminated by ELSTAT if and only if:

a) these data have been treated, as it is specifically set out in the Regulation on Statistical Obligations of the agencies of the Hellenic Statistical System (ELSS), in such a way that their dissemination does not prejudice statistical confidentiality or

b) the statistical unit has given its consent, without any reservations, for the disclosure of data.

The confidential data that are transmitted by ELSS agencies to ELSTAT are used exclusively for statistical purposes and the only persons who have the right to have access to these data are the personnel engaged in this task and appointed by an act of the President of ELSTAT.

ELSTAT may grant researchers conducting statistical analyses for scientific purposes access to data that enable the indirect identification of the statistical units concerned. The access is granted provided the following conditions are satisfied:

a) an appropriate request together with a detailed research proposal in conformity with current scientific standards have been submitted;

b) the research proposal indicates in sufficient detail the set of data to be accessed, the methods of analyzing them, and the time needed for the research;

c) a contract specifying the conditions for access, the obligations of the researchers, the measures for respecting the confidentiality of statistical data and the sanctions in case of breach of these obligations has been signed by the individual researcher, by his/her institution, or by the organization commissioning the research, as the case may be, and by ELSTAT.

Issues referring to the observance of statistical confidentiality are examined by the Statistical Confidentiality Committee (SCC) operating in ELSTAT. The responsibilities of this Committee are to make recommendations to the President of ELSTAT on:

- the level of detail at which statistical data can be disseminated, so as the identification,

either directly or indirectly, of the surveyed statistical unit is not possible;
- the anonymization criteria for the microdata provided to users;
- the granting to researchers access to confidential data for scientific purposes.

The staff of ELSTAT, under any employment status, as well as the temporary survey workers who are employed for the collection of statistical data in statistical surveys conducted by ELSTAT, who acquire access by any means to confidential data, are bound by the principle of confidentiality and must use these data exclusively for the statistical purposes of ELSTAT. After the termination of their term of office, they are not allowed to use these data for any purpose.

Violation of data confidentiality and/or statistical confidentiality by any civil servant or employee of ELSTAT constitutes the disciplinary offence of violation of duty and may be punished with the penalty of final dismissal.

ELSTAT, by its decision, may impose a penalty amounting from ten thousand (10,000) up to two hundred thousand (200,000) euros to anyone who violates the confidentiality of data and/or statistical confidentiality. The penalty is always imposed after the hearing of the defense of the person liable for the breach, depending on the gravity and the repercussions of the violation. Any relapse constitutes an aggravating factor for the assessment of the administrative sanction.

12. Statistical processing	

Detailed information concerning the sampling frame, sampling design,sampling units,sampling size,weights and mode of data collection can be found in the specific section.Such information is mainly used for the computation of the accuracy measures.

12.1. Source data

	Select from here	Answers
A) Survey name in the national language		Έρευνα Υγείας
B) Survey name in English		Health Survey
C) Survey website		http://www.statistics.gr/statistics/-/publication/SHE22/2014
D1) Did the EHIS form part of / was combined with another survey/questionnaire? (This does not include cases when EHIS is implemented as a follow-up of another survey, i.e. when it only uses the respondents but without collecting data of that survey. If EHIS is implemented as a follow-up of another survey it can be reported under other comments.)	YES NO	NO
D2) If YES: The type of the survey(s) that hosted the EHIS questionnaire (multiple answers are possible)	-Health Interview Survey -Health Examination Survey -Disability survey -Living	

	conditions survey -Multipurpose survey -Other, please specify:	
E) Other comments (for example a note in case EHIS is not a part of another survey but is a follow-up of another survey)		

Sampling frame	Select from here	Answers
A1) Sampling frame name:		The selection of Primary Sampling Units (one or more unified city blocks), in the 1st stage, is based on the **General Population Census of the year 2011**.
A2) Data source used for building the sampling frame:	-Population register -Household register -Dwelling register -Population census -List of phone numbers -Postcode address file -Another survey sample (if Yes, indicate which one) -Other, *please specify*	EHIS survey is based on the multiple stage stratified sampling of households from a frame of sampling which has been created on the basis of the results of the **2011 population census** and covers completely the reference population. In order to avoid high non – response rates, before the survey conduct the sampling frames of households are updated in order to have the latest accurated data for all surveyed areas.
A3) Short description of the data source used for building the sampling frame and information included in the data source used for sampling		
B) Frequency of the update procedures of data source used for building the sampling frame	Continuously / Yearly / Quarterly / Monthly / Every __ years / Irregular / None	In each selected Primary Sampling Unit, an updated list of dwellings is compiled for the selection of the dwellings (2nd stage). As the secondary sampling units are the households, all households in each selected dwelling constitute the sample for the data collection. The update procedure of

		the sampling frame takes place before selecting the secondary sampling units (households)
C) Date(s) (or at least the year) of the data source used for the selection of the sampling units for the survey		2011 (General Population Census of the year 2011).

Survey population	Answers
A) Territorial coverage (if applicable): Indicate the parts of the country that were not included	**All the Greek territory**
B) Give an estimate (in %) of the resulting percentage of under-coverage (a proportion of units not accessible via the sampling frame that belong to the target population) (The numerator includes units not accessible via the sampling frame that belong to the target population and the denominator all target population.)	Based on the Census 2011, the estimated undercoverage of the target population is 1.46%, referring to individuals in collective living quarters
C) Give an estimate (in %) of the resulting percentage of over-coverage (a proportion of units accessible via the sampling frame that do not belong to the target population) (The numerator includes out-of-scope units and units of unknown eligibility in the frame population and the denominator all frame population. Duplicates are included in over-coverage. Design-weighted over-coverage rate is preferred to un-weighted rate (the Member State should indicate which indicator was calculated).	0%
D) Population groups that may be excluded even though they belong to sampling frame (e.g. people with hearing problems in case of CATI)	0%
E) Non-target population Give the approximate number of individuals outside the scope of the survey, i.e. the difference between the total country population and the target population, out of which:	All persons aged 15+ residing in the whole territory of Greece at the time of the data collection were included in the survey (one from each household). Persons living in collective households and in institutions are excluded from the target population (non-coverage rate=1.3%)
(E1) population (number of individuals) younger than 15 years (and of whom living in institutions; please define "institutions")	not available
(E2) population (number of individuals) living in institutions (please define "institutions").	not available
E3) If those numbers are not directly available, please provide an estimate. If no estimate is available, please indicate why.	not available

Sampling design	Select from here	Answers
A) Sampling unit(s) (list more options if applicable)	-Dwellings -Households -Individuals	The sampling units are the private households and the individuals. One individual is surveyed from each household of the sample. The target population of individuals is consisted from the individuals aged 15 years or over at the whole country.
B) For dwelling samples Number of households belonging to a selected dwelling interviewed	1 If more than one, how many: All	1
C) For household samples Number of individuals belonging to a selected household interviewed	1 If more than one, how many: All	1
D) Sampling design	Combination of designs (*if so please list the designs used*) Simple Random Sampling (SRS) Systematic Sampling Stratified Sampling (if so, *please indicate stratificationvariables, the categories of those variables and the final number of strata*). Cluster Sampling Multiple Stage Sampling (if so, *please specify the different stages*)	Multiple Stage Sampling
E) Probability used to draw the sample	General comments	**PSU** were one or more unified city blocks. **SSU** were the households of the selected PSU **TSU** were one individual aged 15+ from the selected households
E1) Stage 1: (PSU were …)	equal probability unequal probabilities probabilities proportional to size	Probabilities proportional to size
E2) Stage 2: (SSU were …)	equal probability unequal probabilities probabilities proportional to size	equal probability
E3) Stage 3: (TSU	equal probability	

were …)	unequal probabilities probabilities proportional to size	equal probability
E4) Other stages…	(In case of more complicated sample design, describe it verbally)	not applicable
F) Specific populations oversampled	No Yes, *please specify which ones:*	No
G) Were stratified oversampling methods used?	No Yes, *please specify the reason, what population groups, the method used:*	No
H) Short description of the methods used for drawing up the sample		**1st stage sampling** In this stage, from any ultimate stratum (crossing of Region with the degree of urbanization), the primary units were drawn. The number of draws is approximately proportional to the population size of the stratum (number of households in the last General Population Census of the year 2011). Each area unit (primary unit) in the ultimate stratum has probability of being selected proportional to its size (number of households-according to the 2011 General Population Census) **2nd stage of sampling** In this stage, from each primary sampling unit (selected area) the sample of secondary units (households) was selected. Actually, in the second stage we drew a systematic sample of dwellings. However, in most cases, one household corresponds to each dwelling. The sampling frame containing the secondary units (households) in the selected sampling primary units was updated before the selection of households. **3rd stage of sampling** In this stage, from each selected household, one individual belonging to the target population was selected randomly with equal probabilities.
I) The method used for determining the sample size (what was taken into		For the calculation of the sample size, the minimum effective sample size for Greece (according to Regulation (EC) No. 141/2013) was set equal to 6.667 which was increased by the

account and from where: assumptions on non-response and design effect)		design effect (set equal to 1.5): achieved sample size = 6,667*1.5=10,000	
J) Additional measures taken at the time of drawing the sample to improve representativeness (e.g. choosing the right stratification variables - those that are highly correlated with the variable of interest, etc.)		Stratification of the Primary Units by allocating them according to the degree of urbanization (urban, and rural regions). Except for the two former Major City Agglomerations (Athens and Thessaloniki), which were allocated to 47 nd 11 respective urban strata based on socio-economic criteria. The produced strata according to the degree of urbanization are: Urban Regions with 50.000 inhabitants or more Urban Regions with 30.000 - 49.999 inhabitants Urban Regions 10.000 - 29.999 inhabitants Urban Regions 5.000 - 9.999 inhabitants Urban Regions 2.000 - 4.999 inhabitants Rural Regions 1.000 - 1.999 inhabitants Rural Regions 500 - 999 inhabitants Rural Regions 1 - 499 inhabitants	

	Summary table on participation and non-participation	Number of households	Number of individuals
1	**Total released sample cases** The number of sampling units initially selected from the sampling frame. ► [1] = [2] + [3] + [6]	9936	9936
2	**Ineligible sample cases / out-of-scope units** The unit does not belong to the population of interest for the survey although it is included in the sampling frame.	62	62
2.1	**Non-existent units** The unit does not exist although it was included in the frame due to errors (house/building not existing, no one living in the building/on the address).	21	21
2.2	**Changes in status** The unit has changed its status becoming out of scope for the survey (e.g. change of residence for a household, selected individual died between the reference data of the sampling frame and the moment of the interview, etc.).	41	41
2.3	**Out of target units**	0	0

	The unit has never been in-scope although it was included in the frame due to an inclusion error.		
2.4	**Other ineligible** Other types of ineligibility encountered. It should be specified what are the reasons for this kind of ineligibility: …..	0	0
3	**Eligible sample cases / in-scope units** The unit belongs to the population of interest for the survey (both non-response and response cases). ► [3] = [4] + [5]	9479	9479
4	**Non-response cases / Non-participation** Units for which it has not been able to obtain information.	1256	1256
4.1	**Non-contact** A unit which has not been possible to contact (e.g. no one was at home or wrong address).	524	524
4.2	**Refusal** e.g. selected household or individual was contacted but refused to take part in the survey.	682	682
4.3	**Inability to respond** e.g. selected household or individual was unable to participate due to language barriers or cognitive or physical incapacity to respond (and no proxy interview was conducted).	50	50
4.4	**Rejected interviews** e.g. the selected household/individual did take part but the survey form cannot be used (poor quality - e.g. strong inconsistencies; unacceptable item-response – e.g. individual left most of the questions unanswered; survey form got lost and interview cannot be repeated; etc.).	0	0
4.5	**Other non-response** Other types of non-response encountered. It should be specified what are the reasons for this kind of non-response: …………….	0	0
5	**Response cases / Participation** Units for which it has been possible to obtain information.	8223	8223
5.1	**A fully completed interview** All relevant questions answered by the respondent or by a proxy interview. (a) Variables not recommended to be completed by proxy respondents if the interview was conducted with a proxy respondent and (b) variables for which derogations were granted, are not taken into account when evaluating completeness.	8223	8223
5.2	**A partly completed interview** Not all relevant but at least some technical variables (PID, HHID, PRIMSTRAT, PSU, WGT, PROXY, REFYEAR, REFMONTH, INTMETHOD), sex and age and at least 50 % of all other variables (to be answered and) answered by the respondent or by a proxy interview.	0	0

6	**Unknown eligibility** Selected units with unresolved eligibility.	395	395
	Comments related to the table		

Structure of target, sample, response and non-response population	Target population[1]		Sample[2]		Response		Non-response	
	Number	%	Number	%	Number	%	Number	%
All	**9150412**	**100.0**	**8223**	**100.0**				
Women	4782008	52.3	4827	58.7	0		0	
Men	4368404	47.7	3396	41.3	0		0	
Age (women)	**4782008**	**100.0**	**4827**	**100.0**				
15-24	538737	11.3	284	5.9	0		0	
25-34	673213	14.1	528	10.9	0		0	
35-44	818908	17.1	759	15.7	0		0	
45-54	803967	16.8	796	16.5	0		0	
55-64	699291	14.6	817	16.9	0		0	
65-74	580041	12.1	746	15.5	0		0	
75-84	511077	10.7	684	14.2	0		0	
85+	156774	3.3	213	4.4	0		0	
Age (men)	**4368404**	**100.0**	**3396**	**100.0**				
15-24	524287	12.0	295	8.7	0		0	
25-34	664266	15.2	401	11.8	0		0	
35-44	803022	18.4	552	16.3	0		0	
45-54	743761	17.1	529	15.6	0		0	
55-64	637506	14.6	563	16.6	0		0	
65-74	507804	11.6	518	15.2	0		0	
75-84	390097	8.9	423	12.4	0		0	
85+	97661	2.2	115	3.4	0		0	
Educational attainment level (ISCED 2011)	**9150412**	**100.0**	**8223**	**100.0**				
ISCED 0-2	3828333	41.8	3865	47.0	0		0	
ISCED 3-4	3203169	35.0	2587	31.5	0		0	
ISCED 5-8	2118910	23.2	1771	21.5	0		0	
Labour status	**9150412**	**100.0**	**8223**	**100.0**				
Employed	3309974	36.2	2655	32.3	0		0	
Unemployed	1343764	14.6	993	12.1	0		0	
Retired	2360497	25.8	2808	34.1	0		0	
Students	792637	8.7	458	5.6	0		0	
Other inactive	1343540	14.7	1309	15.9	0		0	

Degree of urbanisation[3]	9150412	100.0	8223	100.0				
Densely-populated area	3884011	42.4	3259	39.6	0		0	
Intermediate-populated area	2595780	28.4	2251	27.4	0		0	
Thinly-populated area	2670621	29.2	2713	33.0	0		0	
Household size (number of persons living in household)	9150412	100.0	8223	100.0				
1	1150044	12.6	2198	26.7	0		0	
2	2481090	27.1	2767	33.7	0		0	
3	2251012	24.6	1490	18.1	0		0	
4	1997772	21.8	1282	15.6	0		0	
5+	1270495	13.9	486	5.9	0		0	
Comments related to the table								

Representativeness of results (National criteria/practices for publication are to be used)	Select from here	Answers

A) For which age groups, resp. combination of sex and age groups, the sample design allows publication of data? (for basic indicators based on HS1, HS2 and HS3 variables) (list all respective options from the list)	- 5-year age groups (Please specify the last age group for which it is possible:)
	- 10-year age groups (Please specify the last age group for which it is possible:)
	- sex and 5-year age groups (Please specify the last age group for which it is possible:)
	- sex and 10-year age groups (Please specify the last age group for which it is possible:)
		5-year age-groups (CV<10%):
		-**HS1**: all age-groups
		-**HS2**: age groups above 30 years
		-**HS3**: age group above 65 years
		
		10-year age-groups (CV<10%):
		-**HS1**: all age-groups
		-**HS2**: age-groups above 25 years
		-**HS3**: age group above 65 years
		
		Sex & 5-year age-groups (CV<10%):
		-**HS1**: male & females in all age-groups
		-**HS2**: males in age groups above 45 years
		      females in age groups above 35 years
		-**HS3**: male and females in age group above 75 years
		
		Sex & 10-year age-groups (CV<10%):
		-**HS1**: male & females in all age-groups
		-**HS2**: males in age groups above 35 years
		      females in age groups above 25 years
		-**HS3**: male in age group above 75 years
		      females in age group above 65 years

B) For which regional breakdown, resp. combination of sex and region, the sample design allows publication of data? (list all respective options from the list)	- NUTS 1 - sex and NUTS 1 - NUTS 2 - sex and NUTS 2 - NUTS 3 - sex and NUTS 3	**NUTSI _(CV<10%)_:** -**HS1**: all NUTSI regions -**HS2**: all NUTSI regions -**HS3**: all NUSI regions -**HO1_1**: all NUTSI regions **Sex & NUTSI region _(CV<10%)_:** -**HS1**: males and females for all NUTSI regions -**HS2**: males and females for all NUTSI regions **HS3**: females for NUTSI regions GR1, GR2 and GR3 **NUTSII _(CV<10%)_:** -**HS1**: all NUTSII regions **Sex & NUTSII regions _(CV<10%)_:** -**HS1**: males and females for all NUTSII regions
C) Does the sample design allow publication of data for your country according to other breakdowns? If Yes, please specify for which breakdowns:		

12.2. Frequency of data collection

ELSTAT collects EHIS data every five (5) years.

12.3. Data collection

Data collection method	Select from here	Answers
A) Data collection method	Unimode Mixed mode	Uni-mode
B) Data collection mode(s) (in case of mixedmode list all the modes used)	- Postal, nonelectronic version - Postal, electronic version (email) - Face-to-face, nonelectronic version - Face-to-face, electronic version - Telephone, nonelectronic version - Telephone, electronic	- Face-to-face, nonelectronic version The paper assisted personal Interviewing (PAPI) technique has been used.

	version - Use of internet - Other (Please specify: …)	
C) Which of the following topics (submodules/variables) were administered via a self-completion questionnaire?	Smoking (SK) Alcohol consumption (AL) other, please list (use EHIS submodule codes that is HS, CD, etc.):	None
D1) Were any variables (including technical and core social variables) completed from a different source (e.g. administrative register)?	No Yes	No
D2) If Yes, please specify which variables and from which sources		
D3) If Yes, why were different sources of data used? (Please describe here possible impact of the use of these other sources on quality)		
Interviewers		
A) Interviewers qualifications	internal staff with experience in health/social surveys external staff with experience in health/social surveys other, please describe: …	For the data collection were used external collaborators who attended one day training course (the duration of each course was approximately 8- 10 hours). Training courses were organised in all Regional offices. Additionally, internal staff with experience in health/social surveys was used (approximately 3 persons)in the data collection phase. External interviewers' selection was made through the ELSTAT's central system for selecting external collaborators. Due to limited human resources, ELSTAT is obliged to use temporary survey workers (interviewers) that do not belong to the permanent staff. The interviewers are paid for the duly completed questionnaires, resulting thus in lower survey cost. In most statistical surveys, the data collection is subcontracted by ELSTAT to temporary survey workers. There is an electronic register for all temporary survey workers, update every 8 months and a ranking system is elaborated. The

		temporary survey workers are selected though transparent, standardized procedures and according to meritocratic criteria.
B) Ratio interviews / interviewers		The number of questionnaires assigned varies from 8 up to 55 questionnaires per interviewer and depends on the available sample in its region and on a maximul number of questionnaires set by ELSTAT, each interviewer can undertake.
C) Describe the interviewer training method and support (e.g. skills testing before starting the fieldwork, duration of training, training materials provided)		The training seminars were carried out by the surveys' project managers and focus both on the survey's concepts and techniques for the data collection as well as on the data entry of collected data in the electronic questionnaires. Emphasis is given on how to handle non-response (refusals).
Quality control during the fieldwork		
A) Method used for the prenotification or first contact of respondents (multiple answers are possible)	letter telephone, incl. mobile personal contact at doorstep internet/email other, *please specify*	An informative (Introductory letter) letter was sent almost one month before the survey conduct in order the surveyed households to be informed that had been selected for the survey and asking for their cooperation and the reliable information provision as well.
B) Use of incentives	No Yes, *please specify*	No use of incentives
C) Theoretical minimal number of contacts with a respondent before declaring a nonparticipation (i.e. the number of attempts that an interviewer is asked to do for interviewing a respondent)		Three call backs/ attempts were asked to the interviewers in case the respondent was temporary absent.
D) Effective (mean) number of contacts really performed before declaring a total nonparticipation (if available)		Information not available.
E) Average, minimum and maximum interview duration for the EHIS questions (in case is not directly possible to indicate, please give an estimate; in case different data collection methods were used, please indicate separately for each of	Average: … minutes Minimum: … minutes Maximum: … minutes	Even though the duration of the interview was not recorded according to the interviewers' reports its interview lasted approximately 45 minutes.

them)		
F) Please describe the techniques used to control interviewer performance		Interviewers' performance control was done via continuous contacts with the interviewers. Thus, we were aware of the progress of the work as well as any problems encountered.
G) Ratio interviewers / field supervisors		In the central service due to the large number of interviewers the monitoring of interviewers' work was allocated into three supervisors. In the regional offices one supervison monitored the work of all interviewers.
H1) Were (some of) the respondents contacted for quality control?	No Yes	No
H2) If Yes, please describe the method (mode of contact, % of respondents contacted, what was checked)		
I) Were additional studies performed in relation to the nonparticipation?	No Yes	No

Questionnaire	Select from here	Answers
Development of the questionnaire		
A) Language(s) in which the survey was carried out		The language used is Greek. In some Regional Units as Rodopi and Xanthi the survey was carried out by using translators due to the use of other languages such as Turkish and Pomakika (a local language of the regions).
B1) Was the translation protocol proposed by Eurostat used for all national languages? (as described in the EHIS Manual, Translation protocol)	Yes No	Yes. The proposed (by Eurostat) translation protocol has been used as described in the Methodological Manual.
B2) If no, please provide a brief description of protocol used (for each language) in the translation process of the EHIS modules		
C1) Were the modules pre-tested?	No (for none of the languages) Yes but not for each language Yes (for all languages)	No
C2) If Yes, please indicate the	simple testing	

type of pre-test (multiple answers are possible) (Note: If 'Yes but not for each language' or if 'Yes' (for all languages) but differently for any language, please describe the situation for each language separately under the item 'other, please specify'.)	cognitive testing behaviour coding special probing expert panel other, *please specify:*	
C3) If 'Yes (for all languages)' or 'Yes but not for each language', did the pre-test concern all EHIS modules/sub-modules or just some of them? (Note: If the situation was different for different languages, please describe it separately)	All of them Not all (please name – using the abbreviations of - modules and sub-modules that were pre-tested):	
D1) Was a pilot field test carried out?	Yes No	No
D2) If Yes, please describe the methodology and organisation of the field test (refer to the objectives of the testing, geographical area covered, age range, sample size, respondents selecting procedure, overrepresented population groups, data collection procedure)		
D3) If Yes, please describe the main lessons learned from the field test		
Design of the questionnaire		
A1) Was the recommended order of modules and sub-modules followed?	Yes No	No
A2) If No, please describe the deviations from the recommended order.		The questions on INCOME have been put in the end of the questionnaire because data is sensitive and has been observed unwillingness on behalf of interviewees in answering questions on income.
B1) Please indicate for which EHIS variables the questions were modified in comparison with the model questionnaire and conceptual guidelines.		Modifications existing concern different wording, added subcategories or totally new questions. In the below mentioned questions there was a slight modification in comparison with the model questionnaire and conceptual guidelines: AC2/ PL1 / PC2 /HA2/MD2/

		PA1/SK1/SK4. The modification of the questions did not affect the requested by Eurostat variables. Questions PA9 / PA10/ FV5/ FV6/ SK3.1 etc (see Annex 3), were totally new questions included in order to cover national needs.
B2) If there were such modifications introduced, please provide here a general description of the changes (splitting into more questions, etc.) and the reason(s) for modifications. If applicable list in Annex 3 detailed description of the modifications and the wording of the modified questions in English		In Annex 3 there are listed all modifications and additions of new questions which took place. The main reason for modifications and additions were national needs. Before the data collection, a working group consisting of ELSTAT employees and health experts from competent ministries / universities etc. was established. In these meetings exchange of views/ clarifications in some questions and the addition of extra questions for national needs were discussed.
C) Please specify the content of alcohol of the "national standard" drink (in grams) used in AL sub-module if different from the guidelines.		The alcohol consumption was counted in glasses in which usually every drink (e.g. glass of beer/wine/whiskey etc) is served. Examples were provided concerning the capacity in mls for each "standard drink " per type, e.g. 1 glass of beer is equal to 250 ml, 1 glass of local drinks as raki/ouzo/tsikoudia is equal to 50 ml . So, on the base of the capacity, if someone reported drinking raki (local spirituous drink) in a full glass of beer we recorded fivefold quantity in glasses.
D) Problematic questions Please indicate what were the modules and/or questions which caused problems during preparation of the questionnaire and/or later during the interviews and which were not detected during the tests. Please describe the problems and ways you managed with them.		The modification we made in SK4 Question by splitting smoking indoors into three cases : A smoking indoors at work / B smoking indoors at home / C smoking indoors at public places, restaurants etc. and by splitting more, frequency of hour exposure (Never or almost never, Less than 1 hour per day, 1-5 hours per day, more than 5 hours per day) caused classification problems in the specific ESTAT variable. The problem was overcome by using specific combination of answers. The combination is available if needed.
E) Core social variables Please comment on implementing		In order to compile variable **HHINCOME** (Net monthly equivalised income of the

the core social variables: possible problems and measures taken to overcome them.		household) : (1) We calculated average net monthly income for each household (for 2756 individuals/households having provided exact income amount we kept that amount, for the rest 5467 cases having provided income in income groups we estimated an average income value for each group (that is, 195 €, 581 €, 933 €, 1328 €, 2241 € for the first 5 income groups of the questionnaire and 3697 € for the last open income group –calculated from the sum of exact income > 2919 provided, divided by the number of individuals having provided it). (2) We calculated equivalised household size for household members (1, 0.5, 0.3) according to ESTAT's manual. (3)For each household net monthly equivalised income was calculated by dividing average total income by equivalised household size. The data were ordered according to the value of the equivalised total net monthly income of the household and the four cut point values were estimated.

12.4. Data validation

Data validation is done by

- conducting qualitative and quantitative tests and
- by using calculation of sampling errors (as a criterion)for the final validation of data

12.5. Data compilation

Describe the data compilation process (e.g. data editing, imputation, weighting, adjustment for non-response, calibration, modelling etc.).	

Item imputation	Select from here	Answers
A1) Did you apply any method to correct for item non-response?	NO YES	NO method was applied to correct for item non-response
A2) If Yes, which method:	- Simple imputation (deterministic) method - Simple imputation (stochastic)	

	method - Multiple imputation approach - Other	
A3) If Yes, please give a short description of the method used (mentioning which auxiliary information or stratification was used):		
A4) If Yes, please provide the values of imputation rate (un-weighted) for all variables for which imputation was carried out: Imputation rate = [number of units for which variable is imputed] / [number of units for which variable is imputed + number of units for which the non-missing value of variable is kept]		

Calculation of weighting factors	Answers
A) Please describe the calculation method in general and underneath in specific	Initially weighting factors for the households (including non-response treatment and calibration) were calculated. Subsequently, each individual received as initial weight the respective household's weighting factor, which was adjusted according to the individual's inclusion probability and calibrated so that the final individuals' weights conform to the population distribution.
A1) Steps: unit's probability of selection, correct non-response, over-sampling/under-sampling of certain population groups, adjust/calibrate the sample to external data relating to distribution of persons in the target population;	A three stage sampling design was used with Primary sampling units one or more unified blocks, secondary unit the household and tertiary unit the individual. Each primary sampling unit (one or more unified city blocks) was selected with probability proportional to size. Each secondary sampling unit (household) was selected with equal probability using systematic sampling. Each tertiary sampling unit (individual) was selected randomly with equal probability within the members aged 15+ of the household's members The extrapolation factor of **individuals** is a product of three factors: a. Inverse of inclusion probabilities

	b. Inverse of response rates in each stratum A factor, which adjusts the sample weights of individuals at NUTSI level, so that the sample distribution conforms to the population distribution across a set of classes. The classes are 14, which are defined by crossing sex with age groups (2 sex categories 7 age groups). The age groups are defined by the year intervals: 15-24, 25-34, 35-44, 45-54, 55-64, 65-74 and 75+. The population distribution of individuals by sex and age groups is calculated using data from the Population Census 2011 and the estimated population for the year 2014. **Household level:** The extrapolation factor of households is a product of three factors: a. Inverse of inclusion probabilities b. Inverse of response rates in each stratum A factor, which adjusts the sample weights of households at NUTSI level so that the sample totals conform to the population totals on a cell-by-cell basis (Population Weighting Adjustment). The auxiliary variable used at household level is the household size (1,2,3 and 4+ members) for the definition of cells or classes. The distribution of households by size class was based on data coming from the Population Census 2011 and the estimated population of year 2014.	
A2) Levels: individual, household, dwelling		Please see A and A1, above.
B) Non-response adjustments (Please describe the method and factors used)	Method: Factors used:	Inverse of response rates in each stratum.
C) Adjustments to external data (including correction for under-sampling/oversampling) (Please describe the method and factors used)	Method: Factors used: - Gender - Age groups; specify which: - NUTS; specify what level/which groups: - Educational	- Gender - Age groups; specify which: 0-14,15-24, 25-34, 35-44,45-54, 55-64,65-74,75+ - NUTS; specify

	attainment level; specify which: - Household size; specify which groups: - Other; please specify: …	what level/which groups: NUTSI level - Household size; specify which groups: 1,2,3,4+
D) Please describe a method used to adjust for extreme weights (trimming)		The final weights were trimmed itiretively so that the weights'CV, not to exceed 75%
E1) Information on the variability of final weights	- average divided by minimum, - maximum divided by average, - coefficient of variation	Average weight/minimum weight: 21.81 Maximum weight / Average weight: 3,59 CV of weights: 0.725
E2) In case a method to adjust for extreme weights was applied, also information on the variability of the weights before trimming	- average divided by minimum, - maximum divided by average, - coefficient of variation	Average weight/minimum weight: 13.58 Maximum weight / Average weight: 15,67 CV of weights: 0.898

12.6. Adjustment