

Sampling design 2022

EHES ROMANIA

EHES manual

- A minimum of 4000 persons are sampled to be invited in each country
- Each of eight age-gender domains (25-34, 35-44, 45-54 and 55-64 years) should have at least 500 representatives in the sample.
- The sample size calculation is based on a participation rate of 70 percent.

Sampling strategies

- Have a lot of commonality (e.g. total number of patients, or the number of patients per PSU, or the estimated response rate)
- Differ in terms of what is considered a PSU

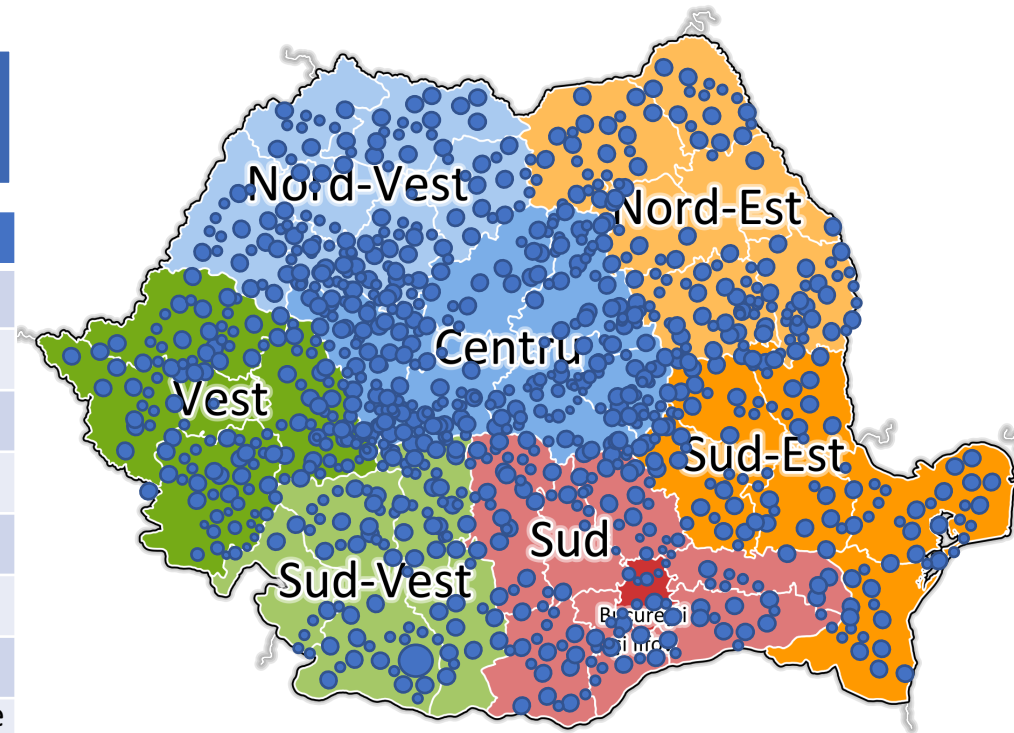
Three main strategies, in the order of their pro/con ratio:

1. Using localities (administrative territorial units)
2. Using electoral sections
3. Using GPs

Sampling strategy #1

National Institute of Statistics (TEMPO database) is the sampling frame available that are proposed to be used in the EHES 2022 survey.

Sampling frame	TEMPO Database
Source	National Institute of Statistics
Last available	January 1 st 2020
Population	all age groups
Population size	Available (by region, county)
Frame count	Administrative Territorial Units (UAT, 3.181 units)
PSU size	various sizes in each strata
Features	Compact, non-overlapping areas
PSU structure	N/A, households selected through a semi-probabilistic method such as the random route
Mapping and listing	N/A
Segments selection	Two stage selection: 1 or more segments per UAT
Estimated sample	9000 gross sample (with an expected participation rate of 50%)
Stratification options	Region, district, locality, locality size, urbanicity



3.181 UATs

UATs as PSUs, multiple sampling points per PSU

Selecting segments in two stage selection - 113 PSUs x 40 HHs

UAT = 3.181 units (compact, non-overlapping)

PSU sampling rate – proportional to the UAT size (PPS systematic)

Stratification by urbanicity and region - $2 \times 8 = 16$ strata

Proportional allocation within strata

Random selection will be used in each sampling stage

The probability of selection for every unit will be retained for analytics

Unique identifiers for every sampling unit selected will be retained for analytics

5200 HHs urban & 3840 HHs rural are randomly selected

Estimated response rate 50% (HH and individual non-response)

At least 4520 questionnaires / medical exams will be achieved, min 40 / PSU

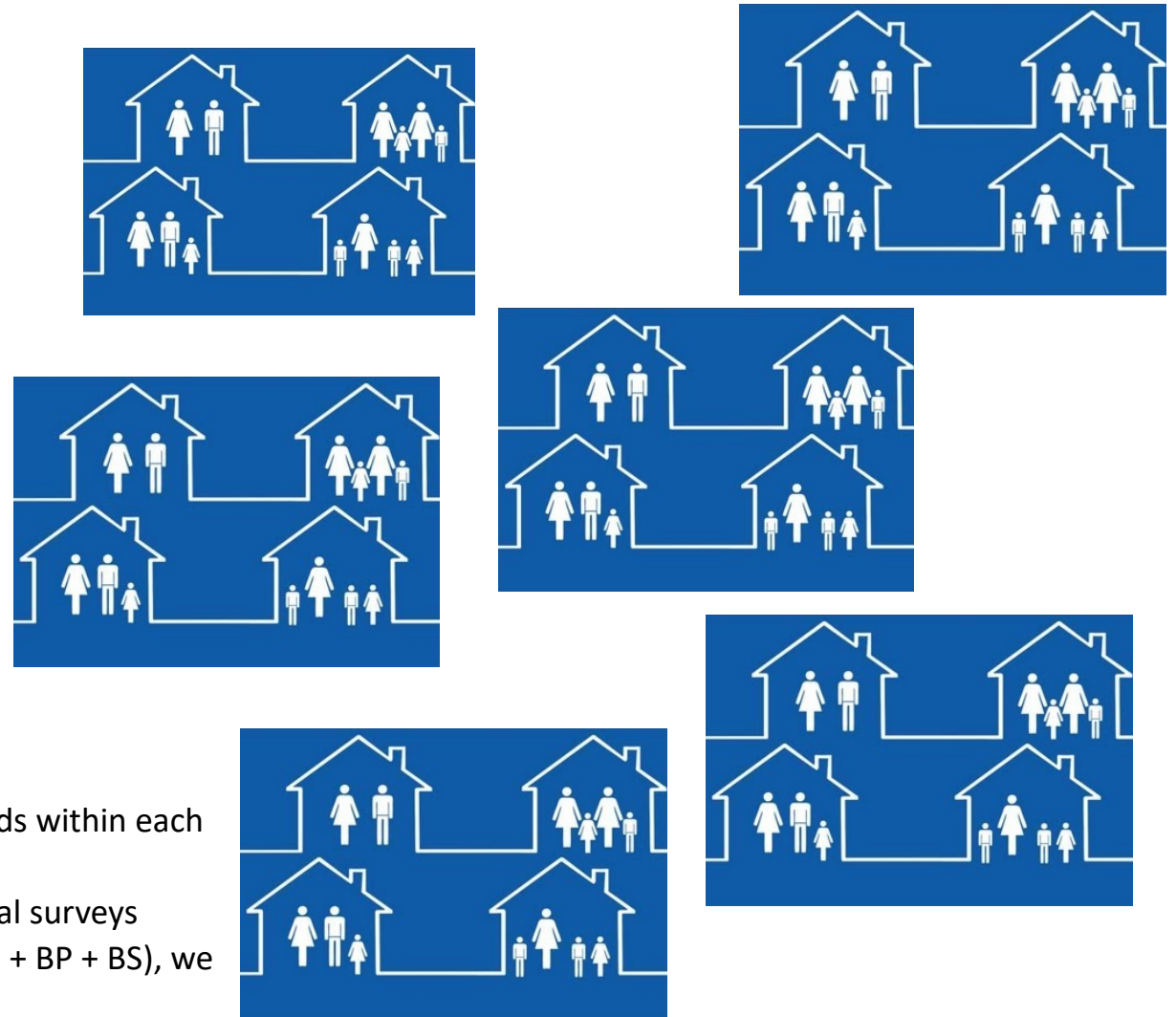
Sampling strategy #1

Number of
eligible
inhabitants

10.608.220	Age group 25-34	Age group 35-44	Age group 45-54	Age group 55-64		
	2.425.586	2.860.569	2.962.503	2.359.562		
	Urban			Rural		
	Number of UAT	Population count	Population %	Number of UAT	Population count	Population %
Bucharest-Ilfov	9	1.252.085	12	32	154.526	1
Center	57	751.515	7	357	509.418	5
North-West	43	795.474	7	403	626.889	6
North-East	46	739.928	7	506	892.166	8
South	48	655.96	6	519	903.394	9
South-East	34	715.694	7	356	558.674	5
South-West	40	524.318	5	408	518.06	5
West	42	637.079	6	281	373.04	4
Total	319	6.072.053	57	2862	4.536.167	43

Sampling frame #1- pros and cons

- recent data (January 1st 2020)
- population statistics available by age groups
- frequently used in Romanian social surveys
- compact and non-overlapping
- the data on the target population is available on EHES age groups

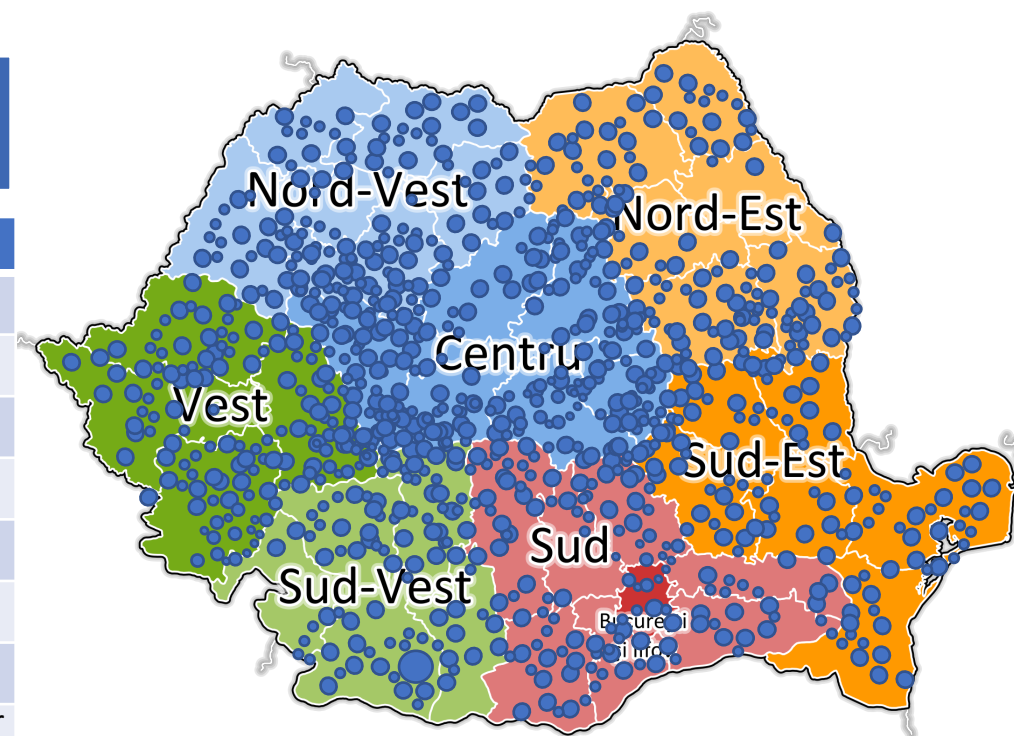


- too expensive to perform mapping and listing of the households within each TAU
- EHES is a more complex (socio-medical) survey than most social surveys
- by implying multiple phases of implementation (questionnaire + BP + BS), we anticipate low response rate
- multiple GPs will be contacted since respondents are heterogeneous

Sampling strategy #2

The polling sections database is the sampling frame available that are proposed to be used in this EHES survey.

Sampling frame	Polling sections Database
Source	Central Electoral Bureau
Last available	November 2020
Population	18+ (<i>with the right to vote</i>)
Population size	Available (by locality, polling section)
Frame count	Polling sections (18.802)
Average PSU size	~950 individuals, ~300 HHs
Features	Compact, non-overlapping areas
PSU structure	Each polling station includes: streets of the locality where it is included or one or more small localities or villages with their addresses.
Mapping and listing	List of streets and numbers available for each polling station
Segments selection	Two stage selection: 1 polling section = 1 PSU
Estimated sample	9000 gross sample (with an expected participation rate of 50%)
Stratification options	Region, district, locality, locality size, urbanicity



18.802 units

**Polling section as PSU
a single sampling point per PSU**

Selecting segments in two stage selection - 113 PSUs x 40 HHs

PSU = polling station = 18802 units (compact, non-overlapping)

PSU sampling rate – proportional to the PSU size (PPS systematic)

Stratification by urbanicity and region - $2 \times 8 = 16$ strata

Proportional allocation within strata

Random selection will be used in each sampling stage

The probability of selection for every unit will be retained for analytics

Unique identifiers for every sampling unit selected will be retained for analytics

5200 HHs urban & 3840 HHs rural are randomly selected

Estimated response rate 50% (HH and individual non-response)

At least 4520 questionnaires / medical exams will be achieved, min 40 / PSU

Sampling strategy #2

*Number of
eligible votes*



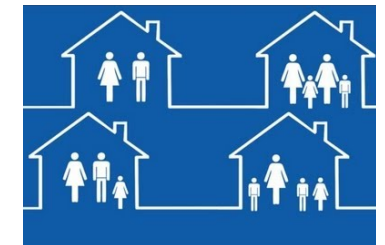
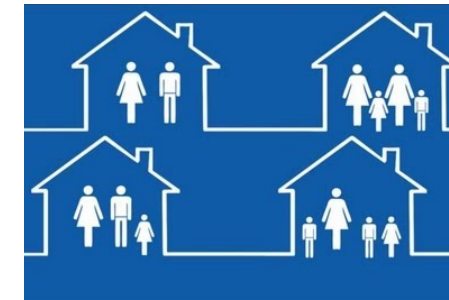
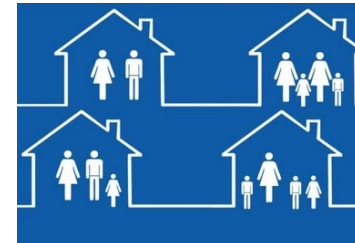
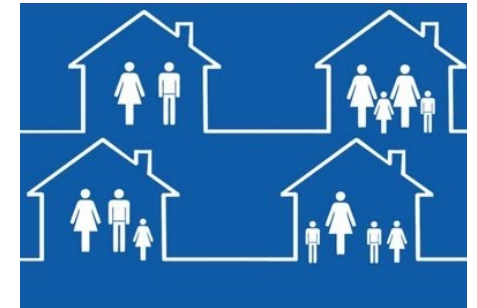
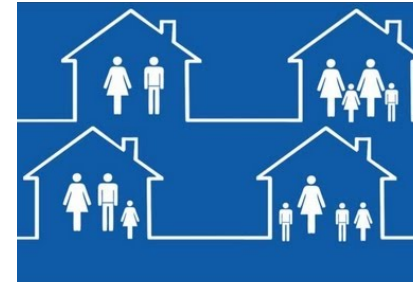
18.216.878	Urban			Rural		
	Number of PS	Population count	Population %	Number of PS	Population count	Population %
Bucharest-Ilfov	1395	1973512	11	154	202453	1
Center	1095	1293450	7	1305	840512	5
North-West	710	861498	5	1355	835643	5
North-East	1207	1522412	8	2285	1723498	9
South	941	1120750	6	1763	1479472	8
South-East	1037	1301492	7	1263	1013030	6
South-West	735	888380	5	1272	891018	5
West	1185	1439856	8	1432	829902	5

Sampling strategy #2 – database excerpt

1	PSU Id	Development region	District	Localitate	Urban/Rural	Populatie 18+	
8766	17771	Bucuresti	MUNICIPIUL BUCUR	BUCUREȘTI SECTORUL 3	Urban	1074	
8767	18418	Bucuresti	MUNICIPIUL BUCUR	BUCUREȘTI SECTORUL 6	Urban	1074	
8768	211	Centru	ALBA	BUCERDEA GRÂNOASĂ	Rural	1073	
8769	21	Centru	ALBA	MUNICIPIUL ALBA IULIA	Urban	1073	
8770	712	Vest	ARAD	GRĂNICERI	Rural	1073	
8771	3097	Nord Est	BOTOȘANI	MUNICIPIUL DOROHOI	Urban	1073	
8772	5420	Nord Vest	CLUJ	MUNICIPIUL TURDA	Urban	1073	
8773	6331	Sud Est	CONSTANȚA	OSTROV	Rural	1073	
8774	9936	Nord Est	IAȘI	BÎRNOVA	Rural	1073	
8775	10049	Nord Est	IAȘI	GRAJDURI	Rural	1073	
8776	12919	Sud	PRAHOVA	ORAȘ BĂICOI	Urban	1073	
8777	16650	Sud Vest	VÂLCEA	MUNICIPIUL RÂMNICU VÂLCEA	Urban	1073	
8778	17351	Sud Est	VRANCEA	VÎNĂTORI	Rural	1073	
8779	6889	Sud	DÂMBOVIȚA	MĂNEȘTI	Rural	1072	
8780	10011	Nord Est	IAȘI	DELENI	Rural	1072	
8781	10029	Nord Est	IAȘI	ERBICENI	Rural	1072	
8782	15334	Vest	TIMIȘ	MUNICIPIUL TIMIȘOARA	Urban	1072	
8783	18230	Bucuresti	MUNICIPIUL BUCUR	BUCUREȘTI SECTORUL 5	Urban	1072	

Sampling strategy #2- pros and cons

- recent data (November 2020)
- population size, list of localities and addresses available for each polling station
- since this is a sample of households, all age groups are covered
- it also covers the foreigners working or living in Romania (non-citizenship) – approx. 0.5%
- gold standard in sampling design in Romania for social surveys
- compact and non-overlapping

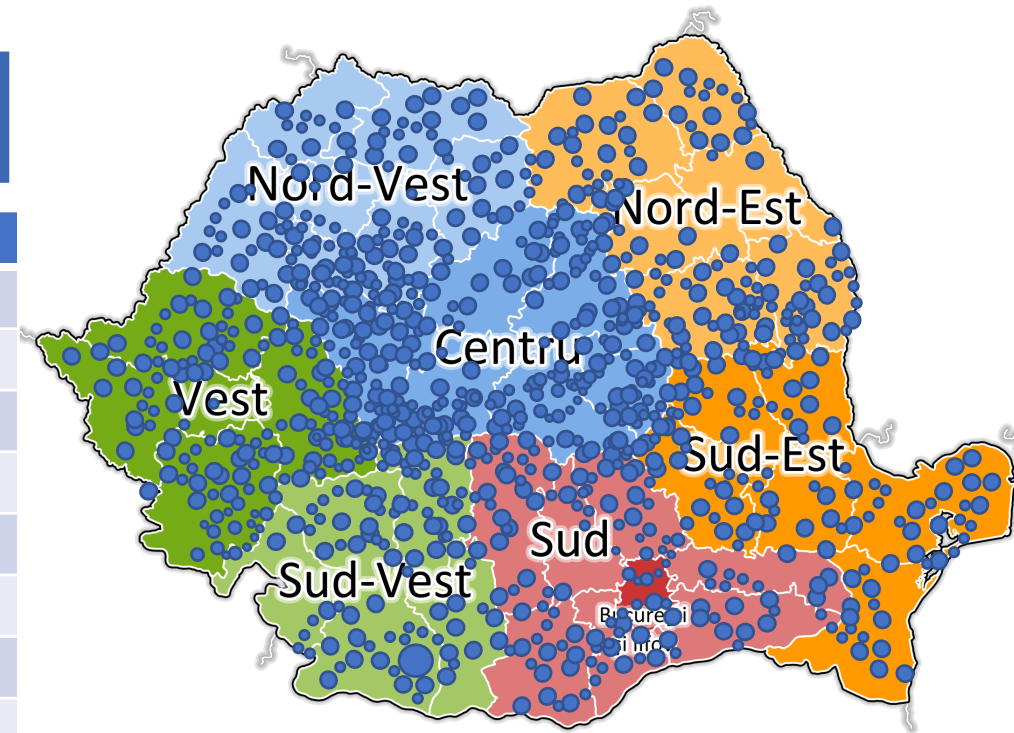


- costs for mapping and listing of the households within each polling station
- EHES is a more complex (socio-medical) survey than most social surveys
- by implying multiple phases of implementation (questionnaire + BP + BS), we anticipate low response rate
- multiple GPs will be contacted since respondents are dispersed at various GPs
- some polling stations have a small size (< 40), but these can be easily redistributed within the same strata to a bigger station

Sampling frame #3

Family Practitioner database is the sampling frame available that are proposed to be used in the EHES 2022 survey.

Sampling frame	Family Practitioner Database
Source	National Health Insurance House
Last available	To be confirmed
Population	All residents with health insurance
Population size	Available (by locality, county, gender, age)
Frame count	Family practitioners (12.187/ NIS TEMPO Database, 2019)
Average PSU size	~1000 patients (to be confirmed)
Features	Non-overlapping units, easy to select
PSU structure	First stage of PSU election is family practitioner, second stage is the potential participant in the study from FPs records
Segments selection	Two stage of selection: 1 GP = 1 PSU = 40 patients (PPS)
Estimated sample	9000 patients (gross sample)
Stratification options	Region, district, locality, locality size, urbanicity, gender, age groups



12.187 general practitioners

Net sample: 40 respondents x 113 GPs

1 GP = 1 PSU

Selecting segments in two stage selection - 113 GPs x 40 patients

PSU = GPs = 12.187 units (non-overlapping)

PSU sampling rate – proportional to the size (number of patients) per GP

Stratification by urbanicity and region - $2 \times 8 = 16$ strata

Proportional allocation within strata

Random selection will be used in each sampling stage

The probability of selection for every unit will be retained for analytics

Unique identifiers for every sampling unit selected will be retained for analytics

The 113 PSUs are proportionally allocated to the urban/rural population

Estimated response rate 50% (individual non-response)

At least 4520 questionnaires / medical exams will be achieved, min 40 / GP

Sampling frame #3- pros and cons

- recent data **(to be requested from CNAS)**
- we estimate a higher participation rate
- lower costs/ GP and lower logistical costs of implementation
- lower costs of selecting potential participants because respondents will be selected from fewer GPs
- suitable sample design for complex socio-medical survey
- non-overlapping
- the support from the GPs



- the study will be national representative only for individuals with health insurance

