4PCAN

D5.1 – PNAs Methodology

2024

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

4PCAN

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

Project number: 101104432
Project name: Personalized CANcer Primary Prevention research through Citizen Participation and digitally enabled social innovation
Project acronym: 4P-CAN
Call: HORIZON-MISS-2022-CANCER-01-01

Version number	
Status	Final
Dissemination level	Public
Due date of deliverable	30/04/2024
Actual submission date	30/04/2024
Project officer	Marianne Ines DA SILVA
Work package	WP5 Living-Labs for primary prevention of cancer
Lead partner	INOMED
Partner(s) contributing	-
Authors	
Main author name	Prof. Gabriel-Marian Hâncean (INOMED)
	Iulian Oană (INOMED)
	Bianca Mihăilă (INOMED)
Contributing authors	Marius Geantă (INOMED)
	Cosmina Cioroboiu (INOMED)
Reviewers	
Reviewer name	Adriana Boată (INOMED)

Statement of originality *This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.*

Version Tracker

Date	Version number*	Change
29/04/2024	0.1	First Draft
30/04/2024	0.2	Final

Table of Contents

List of Figures

List of Tables

List of Abbreviations

- **COVID-19:** Coronavirus Disease 2019
- **HPV:** Human Papillomavirus
- **PNA:** Personal Network Analysis
- **SARS-CoV-2:** Severe Acute Respiratory Syndrome Coronavirus 2
- **SNA:** Social Network Analysis
- **SRFs:** Social Risk Factor(s)
- **UV:** Ultraviolet
- **WHO:** World Health Organization

Executive Summary

Cancer is the second cause of death in all European Union countries, including Romania. The purpose of this deliverable is to introduce the Personal Network Analysis (**PNA**) methodology and show how it can be used for the quantitative analysis of the relationship between social networks and cancer risk factors (overweight, diet, physical activity, smoking).

We describe the context, motivation, study outline, and methods of our study as applied in one of 4P-CAN's living labs, a user-centered open-innovation ecosystem operating in Lereşti – a small rural community of approximately 5,000 people in Argeș county, Romania. This living lab is part of the work carried in WP5 Living-Labs for primary prevention of cancer of the 4P-CAN.

People are interconnected, and therefore their health is interconnected. Yet, there is still limited research on the co-evolution of social networks and cancer risk factors. The current state of knowledge regards people as acting in isolation, and it mostly focuses on individual-level risk factors (like age, sex, and genetic predisposition) and lifestyle choices. Here we present the PNA methodology by describing what social network analysis (SNA) and personal network analysis are, how the network approach can be used for the study of cancer primary prevention, network-oriented research objectives and hypotheses, types of variables specific to PNA, the recruitment process used in the Lereşti living lab, the questionnaire testing phase (i.e., the pilot phase), and further modifications after feedback received through citizens' engagement.

The intended audience consists of persons who are interested in the application of network-oriented methods on topics pertaining to public health. They can be academics focused on innovating the current state of knowledge, or policy-makers who want to address public health and prevention problems beyond the individual level of biological characteristics and health-related behaviours, by looking at how people are influenced in their (health) actions and beliefs by those surrounding them.

We describe the 4P-CAN PNA study, which uses a personal network research design and deploys a prospective cohort study on a sample of approximately 80 participants in Romania. We show the data collection process, focused on information about participants' characteristics, their contacts, and the relationships among these contacts. Three repeated observations (waves) and a pilot study generate a complex data structure analyzed by advanced statistical network models. We expect to detect clusters of cancer risk factors (i.e., people who smoke, are overweight, do physical activity, etc., tend to share social ties). We

also want to assess to what degree these clusters are attributable to social influence, social context and social selection factors. Our study may have implications for primary prevention measures and the efforts of the European Commission to decrease cancer cases by 2040.

Deliverable Introduction

Deliverable objective and scope

The objective of this deliverable is to:

- Present the Personal Network Analysis (PNA) methodology
- Present the need for linking PNA methodology with preventive medicine
- Describe the results of the first waves of the study

Relation to other WPs and deliverables

The present deliverable is linked with the work that will begin in the second year of the project in the Bulgarian Living-lab (the second of 4P-CAN project), providing the framework and the validation of the social network analysis carried out in Plovdiv Region.

Content of the deliverable

The deliverable contains:

- The motivation for using a personal network approach in the study of cancer risk factors.
- An outline of the study undertaken in the Lerești 4P-CAN living lab, describing the research objectives, and methods.
- Possible risks and alternative approaches.
- A presentation of the first two waves of the 4P-CAN Personal Network Analysis study.
- Appendix describing the ego-centric questionnaires used in the pilot phase (wave 1) and in the second wave.

Context

Cancer is the leading cause of death worldwide, with a toll of ten million cases in 2020[1]. Further, approximately four million new cases are reported each year in Europe. On top of that, estimations show that the number of people diagnosed with cancer will increase by up to 18% in 2040[2]. In Romania (a southeastern European country with a population of 19.12 million), cancer is the second mortality cause after cardiovascular diseases, with 98,886 new cases reported in 2020[3].

Some cancer factors are *modifiable*, such as tobacco use, diet and nutrition, physical inactivity and obesity, alcohol consumption, sun exposure pollution, or infections (HPV, Hepatitis B & C). Others are *nonmodifiable*, like age, genetics, sex, personal health history, or immune system suppression (i.e., the body's protection from pathogens, such as bacteria, viruses, and other infectious microorganisms). Recent work[4] claims that approximately 78% of health is affected by social determinants (*what an individual does* – 36%, *how an individual is connected* – 24%, *medical care* – 11%, and *the socio-economic environment* – 7%). It may come as unexpected that *genetic and biological* factors account for only 22% of health. Put differently, managing health's social determinants may prove critical, especially in medical prevention and, indirectly, in decreasing cancer cases and other medical conditions. In this respect, WHO estimates that preventive measures should reduce incidence and mortality while 30 – 50% of all cancer cases are preventable[1].

Despite the existing preventive strategies (e.g., healthy choices, such as quitting smoking, or supra-individual actions, such as tobacco control), various barriers influence the effectiveness of the preventive measures. First, there are *individual barriers*: psychological factors (fear, denial, fatalism, etc.) or difficulties of behavioral change (quitting smoking, eating healthy, etc.). Second, there are *macro-level barriers*: socio-economic disparities, cultural barriers, policy and regulatory changes, inadequate institutional collaboration, etc. Third, there are aspects directly related to *the complexity of the cancer*. We may refer here to genetic factors making prevention more complex, preventive measurements having delayed effects, misinformation, and conflicting medical information.

With very few exceptions, extant studies address individuals in isolation despite that factors related to individual behavior explain more than 70% of health. Disregarding social structures is striking as people are interconnected, and so is their health. Under this context, bringing forth *the network approach*, we may expect to produce new insights for preventive medicine, assess the current

preventive strategies, and test public health interventions and their corresponding effects.

Motivation behind the study

Current limitations and difficulties

The *4P-CAN project*[5] studies the relationship between *social networks* and *the risk factors associated with various types of cancer*. The current state of knowledge related to social risk factors of cancer is mainly based on individual-level risk factors (like age, sex, and genetic predisposition) and lifestyle choices (like smoking, diet, and physical activity)[6]. There is still limited research on how these risk factors spread or are influenced within social networks. Yet, social network analysis can be computationally intensive, especially with large datasets. Additionally, ethical considerations arise when collecting and analyzing personal networks[7] and health data.

Remarkably, prospective cohort studies are scanty and restricted to descriptive network measurements like network size and density. With very few exceptions[8,9], the extant longitudinal research simply neglects advanced network models. Moreover, social network analysis is inherently complex due to the multitude of factors that can affect social relationships and behaviors. Understanding how these complex interactions contribute to health outcomes is a challenge. In conclusion, this research topic is critical from a scientific and socio-economic perspective. While several challenges are associated with social network analysis (high-quality fieldwork data is costly and laborious to collect, and cohorts are hard to maintain), the potential benefits and advancements it could provide are substantial. Mixing social network analysis with medical research is a burgeoning field that promises to improve our understanding and management of cancer.

The elements of originality and innovation related to the current state of knowledge

The mix between *social network analysis* (henceforth SNA) and *medical research* is only at the beginning in many medical areas such as cardiovascular health[10] or cancer research. First, this project brings forth a new understanding of risk factor dynamics. Traditionally, risk factors such as overweight, smoking, physical activity, and diet are viewed from an individual-centric perspective[10]. By employing SNA, we could highlight the dynamics and spread of these risk factors within a social context. Analyzing how social norms and peer influences affect these risk factors can provide a new layer of understanding. Second, utilizing the data obtained from

social networks could develop predictive models that identify individuals or groups at a higher risk of developing cancer. This innovative approach goes beyond simple risk identification and towards preemptive healthcare. Third, the insights gained from this study could lead to developing novel, network-based intervention strategies. Beyond focusing on individual behavior change, these interventions target groups or influential nodes within social networks to encourage widespread health behavior changes.

Fourth, the fusion of SNA with public health research is an innovative, interdisciplinary approach. It opens up new possibilities for collaboration and a broader understanding of health behaviors. Fifth, incorporating modern technologies, like statistical network algorithms, can enhance the analysis of longitudinal complex network data. This way, we can improve the quality of research, provide more accurate results, and handle large datasets. Sixth, the project also considers vulnerable individuals (e.g., low-income, comorbidities, senior citizens) where cancer may be more prevalent. That can bring an element of social justice and health equity into the discussion, which is innovative in its application and important for policy implications. Lastly, this project generates unique longitudinal network datasets on cancer risk factors that will allow replication and secondary data analysis.

Study outline

Research objectives

Traditional public health perspectives analyze individual choices for unhealthy lifestyles (e.g., smoking or processed food intake). As already stated, in the case of cancer, the most salient risk factors have a consistent social component, e.g., lifestyle choices or diet. Therefore, the mitigation of these factors should lead to improvements in cancer prevention. Social (personal) networks are fundamental to how risk factors associated with cancer spread, persist and dissipate.

Social (personal) networks profoundly shape people's views and behaviors. At the same time, people significantly shape their networks. Modeling the co-evolution of networks and behavior is an essential tool to parameterize real-world health models. Thus, controlling for social selection (homophily effects) and other confounding effects (spatiotemporal colocation or social spaces) is critical for identifying the impact of networks (social contagion) on social risk factors like obesity, smoking, or diet.

This study examines *the co-evolution of social networks and cancer risk factors in the context of other variables (e.g., sex, age, education, income), both controlling for these traits and considering interactions with them* - the main research objective of the project.

Given their relevance to cancer, we consider the following social risk factors: smoking, diet (including alcohol consumption and sodium intake), overweight, and (non-related-to-work) physical activity. We henceforth dub these factors SRFs (social risk factors). We consider several specific research objectives of different complexity (**Table 1**). These objectives directly contribute to the project's main aim and rely on panel network data (prospective cohort information).

Specific research objectives	Examples (research network hypotheses)
#1. To determine if individuals with a particular SRF are more likely to have peers displaying the same SRF (Assortativity models) & if SRFs affect network properties.	• People who smoke are more likely to be surrounded by people who smoke. • The networks of people who eat processed foods display different properties from those of people who do not eat processed foods.
#2. To examine whether people who share social ties are similar on several SRFs (Social influence models).	• Friends tend to share the same diet habits (e.g., processed food consumption).
#3. To study whether SRFs predict social relationship patterns (Social selection models).	• People who share the same non-related-to-work physical activities (e.g., jogging) tend to be friends
#4. To test how an SRF and an individual's network position are inter-related (Co-evolutional models).	• Individuals' body overweight (e.g., obesity) covary with their network structure (obese people tend to interact with obese people and reinforce their diet habits).
#5. To analyze if patterns of contacts in temporally explicit data are associated with SRFs (Relational hyperevent models).	• The personal networks of people who smoke have a different pattern and composition in time.
#6. To introduce and evaluate network-oriented public health interventions to control SRFs (Public health interventions).	• Health education and support initiatives through network influencers impact the features of the social networks.

Table 1 Study's specific objectives with examples of network hypotheses

Methods

The present study is based on a social network analysis framework (SNA), whose assumptions about social phenomena are centered around the importance of interrelations[11]. As a research field, SNA offers specialized theories and methods for analysing and explaining the variety of relations and interactions in which social actors (usually persons) are engaged in. Social networks are comprised of *nodes* and *ties*. Nodes can represent various types of social actors. They can be persons, countries or geographic regions, households, companies, social media accounts, etc.[12] Ties, on the other hand, represent relationships or interactions with social and analytical importance between such social actors. For example, ties can show if the nodes have friendship or family relations, if they communicate and how often, or if they attended the same events. Consequently, SNA takes into account two data types: *attribute* and *relational* data[13]. Attributes refer to the particular traits of the nodes (e.g., age, sex or gender, smoking status, subjective health perception, or education). Relational data complement attributes, by adding information about social relations, interactions, or any other relevant data that can be displayed as a tie between two nodes. Compared to other research methods, SNA combines both types of data for inferences about nodes, relations, or the network as a whole.

Given the intertwining of attribute and relational data, networks can be described through composition and structural measures. *Compositional* characteristics are based on nodal attributes – e.g., the proportion of women (or men), average age, or distributions of health behaviours. *Structural* measures are determined by how nodes are connected inside the network. For example, we can characterize networks in terms of their fragmentation (disconnected nodes and clusters), density (cohesiveness level), or degree of centralization (if the network is dominated by a few nodes with many ties).

Social network analysis presents two broad approaches for the design of a network study: *sociocentric/whole network analysis* and *egocentric network analysis*[14]. Sociocentric designs are based on bounded populations of nodes and the ties between them (e.g., friendship relations between all employees inside an organization). On the other hand, egocentric designs look at the networks that are centered around a sample of focal nodes, dubbed *egos*, by investigating the compositional and structural characteristics given by the persons surrounding the ego, named *alters*[14]. In this design, egos report about themselves and about their alters. As such, PNA data can be split into: *ego-level data* (respondent's attributes); *alter-level data* (alters' attributes); *ego-alter data* (various tie characteristics between ego and her alters); *alter-alter data* (various ties between alters).

The specific of a personal network design is that it takes into account the multiple social contexts (and groups) in which a person is embedded. While sociocentric designs and standard egocentric analysis are focused on a single social context (e.g., the workplace or a school), PNA samples alters with various characteristics and importance for the ego (e.g., family members, close friends, co-workers, or acquaintances)[14].

We deploy a *prospective cohort study* (at least three-wave network-panel) with participants recruited from an already established living lab (a user-centered open-innovation ecosystem operating in Lereşti – a small rural community of approximately 5,000 people in Argeş county, Romania). This living lab is part of the 4P-CAN Horizon Europe project (HORIZON-MISS-2022-CANCER-01, Research & Innovation action), focused on personalized cancer primary prevention research.

We use a cohort of at least 80 individuals chosen through ethnographic fieldwork methods and link-tracing sampling. We use variables such as sex, age, and education in recruiting the panel participants. Expectedly, based on the available literature, these recruitment variables should give access to heterogeneous participants in terms of experiences and network features. By

mixing the recruiting variables, we anticipate participants of different profiles in the cohort, e.g., poorly educated young males, poorly educated senior females, highly educated adult females, etc.

We plan to perform a series of (at least) three repeated observations (three waves), i.e., we interview each panel participant thrice. The observation points are roughly equally spaced within a time interval between 18 and 24 months. **Fig. 1** illustrates the research design and provides a glimpse of the empirical data structure. We perform face-to-face interactive interviews collecting the data with Network Canvas.

We ask the panel participants (the egos) to elicit a list of alters based on several *name generators:* 1) *household members*; 2) *other family members and relatives*; 3) *co-workers*; and 4) *friends*. In all cases, the condition for an alter to be taken into account is their *daily or weekly face-to-face interaction with the ego*. For illustrative purposes and readability, we kept only ten alters in **Fig. 1**. Specifically, in **Fig. 1**, the blue node represents a panel participant (Ego 1). All the nodes directly connected to the ego are her social contacts. Further, the interviewees report various data about themselves and their social contacts (e.g., age, sex, education, smoking status, subjective health status, consumption of processed foods, etc.). Next, the participants give information about their relationships with the alters. Lastly, we ask the ego to report, in a pairwise fashion, whether the social contacts *know each other and if they interact at least once a week or less often than weekly*, i.e., in **Fig. 1**, the red lines which mark social contacts that know each other by name and can interact in the absence of the ego. The width of lines indicates the weight of alter-alter ties, i.e. their interaction frequency.

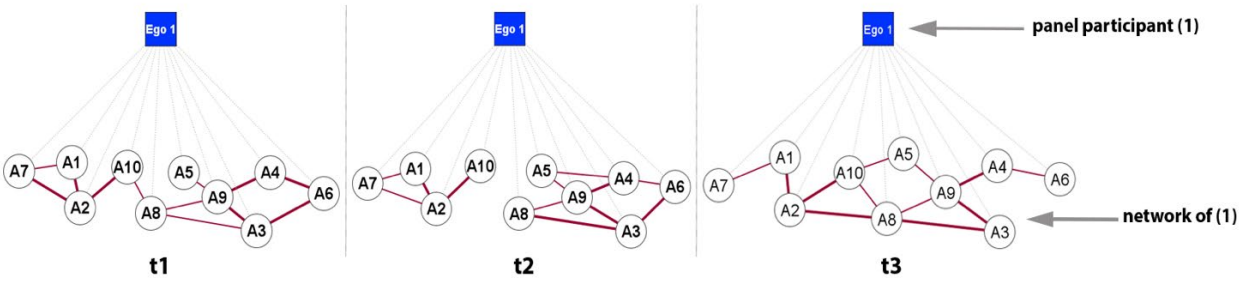

Figure 1 Dynamics of a personal network through different waves

Variables and measurement

Table 2 briefly presents the variables of interest and their measurement level. Of course, the list is not exhaustive, and we might make minor adjustments between

the phases, based on participants' feedback and further data analysis (e.g., an item which presents no variation). We perform actor and network-level measurements. These variables provide basic empirical raw information that underpins testing statistical models and investigation of various social processes (e.g., social contagion or homophily).

The actor-specific variables refer to both interviewees and their social contacts. In the case of alters, the information comes from the ego and is validated afterward (in the subsequent waves). We use the information about socio-demographic features and social risk factors to approximate a cancer risk status for each participant (the vulnerability to cancer).

Variable type	Illustration
#1. Characteristics of the individuals	
Actor-specific	• (socio-demographic data) Sex, age, education, income, marital status, labor market status. • (social risk factors) Tobacco smoking, body overweight, non-related-to-work physical activity, processed food consumption. • cancer risk status.
#2. Network measurements	
Ego-alter ties	• household members, kinship, co-workers, friends.
Alter-alter ties	• if know each other and how often they interact (at least once a week or less often).
Network-level	• density (the observed ties by possible ties), triadic census (how many triad types are present), centralization (the organization of relationships around a specific node), clustering (the detection of densely connected network subparts), modularity (the degree of network division), components (the number of disconnected parts in a network).
Node-level	• degree centrality (the count of social ties that a node has), betweenness centrality (how often a node occurs on all shortest paths between any two other nodes in the network).
#3. Context variables	
Seasonality	• seasons influencing the dietary behaviours of study participants

Table 2 Variables and measurement

The network measurements include structural information about the nodes (ego and alters). At the node level, centrality variables gauge the amount of influence an alter has in a participant's network. As we work with personal networks

(a special type of social network), the node level measurements mostly apply to the alters. The network-level variables assess the characteristics of the personal networks as a whole. We also measure tie properties. Alter-alter ties indicate dichotomous relationships (either present or absent). Ego-alter tie variables refer to whether the ego and her alters live in the same household, are kin/family members, are co-workers, or are friends. The context variables pertain to seasonality which influences the dietary patterns of egos.

Models

We address the project's research objectives in several ways. First, we develop social influence models that use individual traits (e.g., smoking, the cancer risk status) as dependent variables while accounting for the network structure. The social influence models suggest that connectivity breeds influence (social contagion) and, ultimately, explains the behavior. A possible application is the use of stochastic actor-oriented models (*SAOMs*) implemented in *RSiena*[15] for parametrization estimation. Second, we advance statistical models that use networks as an outcome variable (e.g., tie prediction). These models should examine the social processes responsible for generating social structures (e.g., similar people develop ties, homophily). A possible example would be the deployment of exponential random graph modeling (*ERGMs*) implemented in the *statnet* suite[16]. Third, we may consider the data structure presented in **Fig. 1** as *hierarchical*, i.e., alters (level 1) are nested in egos (level 2). In this vein, we can try to fit generalized linear mixed models, as implemented in *lme4* R package[17], to assess assortativity in the participants' networks. (Assortativity marks the tendency of people sharing a tie to be similar on a specific set of attributes.) Fourth, we may build relational hyperevent models (similar to *RHEMs* & *RHOMs*), in *eventnet*, to estimate if patterns of contacts in temporally explicit data are associated with specific SRFs. All these models are also useful for capturing the potential change introduced in some networks by public health interventions.

Scientific feasibility

There are extant findings that support the scientific feasibility of the study. The general scientific aim of the project and the specific research objectives build on the idea that social networks play a significant role in the propagation of risk factors associated with cancer. This working hypothesis relies on the following findings in the field.

Obesity and Dietary Habits. Studies have shown that social networks can influence dietary habits[18] and physical activity levels[19], which are directly tied to obesity[20], a significant risk factor for cancer. For instance, the Framingham Heart Study found that a person's chances of becoming obese increased by 57% if they had a friend who became obese[9].

Smoking. Friends and relatives can significantly influence a person's decision to start or quit smoking[21,22]. The same Framingham Heart Study found that when a person quits smoking, their friends' chances of quitting increased by 36%[8].

Physical Activity. Research has shown that people tend to emulate the exercise habits of those around them[20]. If a person's network is active, they will likely be active, and vice versa[19,20].

Socio-economic Status and Health Inequality. Finally, there's evidence that socio-economic inequalities in health can spread through social networks. People who are socially connected tend to have similar economic status and consequently similar health outcomes, including cancer risks[23].

In sum, our study seeks to extend the extant findings by examining the co-evolution of networks and behavior (cancer risk factors). This examination controls for seasonality and disentangles processes like social selection and influence. Remarkably, the fusion of social network analysis and medical research is only at the beginning in areas such as cancer or cardiovascular health.

Citizens' engagement

A living lab is characterized by the engagement of the local community in the project, as it can be viewed as an open-innovation ecosystem where the participants actively engage in the co-creation of the study methodology by voicing their needs and expectations. During the prospective study, inside the living lab are organized various events at which the study participants are invited. At these events, the citizens are informed about prevention measured in relation to the cancer risk factors – e.g., workshops about healthy eating. More so, the 4P-CAN events are both a way of keeping the participants interested in the study and an opportunity to capture dynamic co-participation data suited for analyzing relational hyperevents models (*RHEMs* and *RHOMs*) or other frameworks using co-event data.

Risks and alternative approaches

Participant Drop-out or Attrition: For a 36-month longitudinal study, there is a risk that participants might drop out, impacting the ability to collect sufficient data for meaningful analysis. As mitigation solutions, we consider regular & engaging communication with the study participants and incentives for continued participation.

Data Collection Issues: Errors or inconsistencies in data collection can lead to flawed results. Training the researchers thoroughly and piloting the data collection instruments can help address this. We will regularly check the collected data for errors or inconsistencies and make necessary adjustments.

Unexpected results or difficulties in analysis: The data might not show clear patterns, or the network analysis might be more complex than anticipated. Therefore, we will make adjustments in methods to adapt to unexpected challenges.

Lack of adoption or impact of the proposed interventions: The interventions proposed based on the study results might not be adopted by the participants or have the desired impact. This risk can be mitigated by involving the participants in the intervention design process (e.g., via citizen engagement, focus groups, or interviews), ensuring the interventions are well-matched to their needs and preferences. Further, the research team will collect insights from the other research teams and stakeholders in the *4P-CAN living lab*.

Privacy and ethical concerns: Given that the project analyzes social network data, there could be potential privacy and data security concerns. Strict protocols were implemented for data handling, storage, and analysis to keep participant information confidential. All the procedures will follow relevant ethical guidelines. Appropriate institutional review boards provided and will continously provide ethical approvals (i.e., University of Bucharest, Centre for Innovation in Medicine and the 4P-CAN Ethics Advisor).

The first wave (the pilot phase)

The pilot phase of the study, i.e., the first wave, took place between June and September 2023, consisting in ethnographic fieldwork and face-to-face interviews. Before delving into the actual data collection phase, meetings were held with local authorities and community members, during which the project team discussed the project's scope and primary objectives. These meetings aimed to familiarize citizens with the project team and instil confidence in the importance of the scientific endeavour. They were disseminated through the local press via mass media. Promotional videos are available on the project's social media pages (LinkedIn, Facebook, and X), as well as on the Facebook pages of the local press, which was invited to these events.

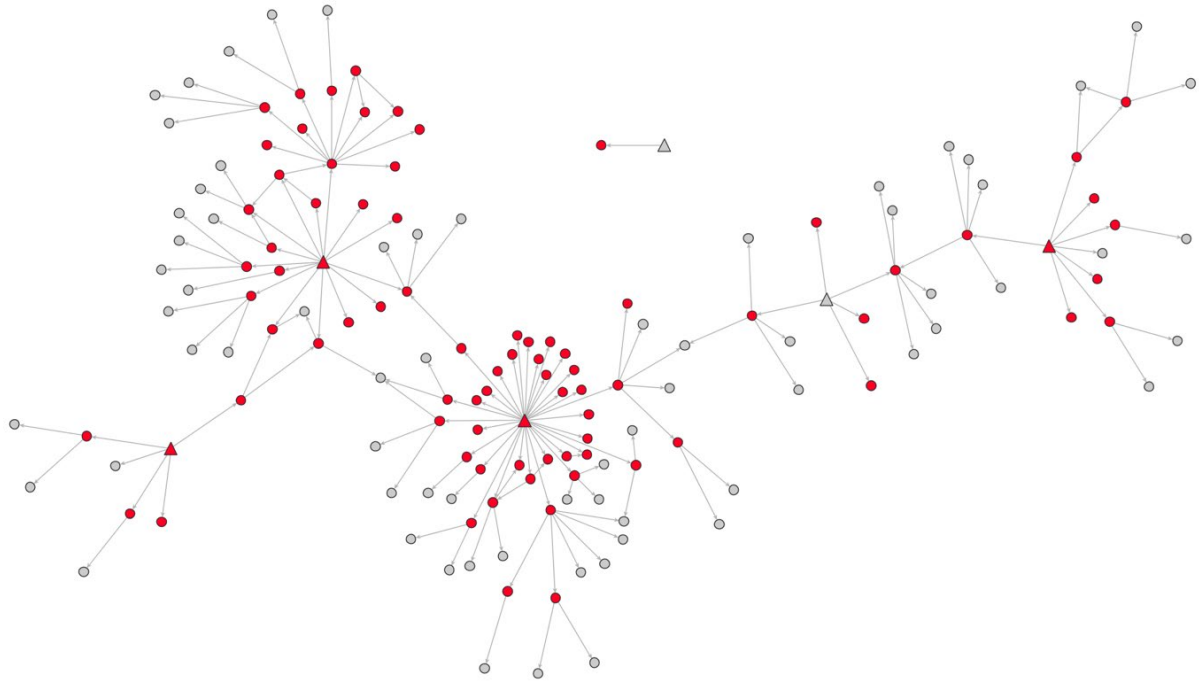

Figure 2 Link-tracing sampling. Referee-referral chains
Note: from Hâncean et al.[24]

As a methodology for participant recruitment, we employed a link-tracing sampling design[25,26]. Building upon previous ethnographic research, held between June and September 2023, we selected 6 seeds, ensuring diversity to encompass various socio-economic indicators. Consequently, our seeds differed in terms of sex (two females and four males), age (ranging between 30 and 64 years old),

education (one person with secondary level studies and five with tertiary level), personal income (half below the average national salary and the other three above), and status on the labour market (one self-employed, one retired, and four employed). Following the interviews with the seeds, we proceeded to gather data from individuals recommended by each seed, and subsequently from those recommended by the initial recommendations, and so forth, building referee-referral chains. **Fig. 2** illustrates the recruitment process, where *triangles* indicate the seeds and *circles* indicate the referrals made by the seeds or by those recommended by the seeds and so on. *Red* nodes show the persons who participated to the interview, while the *grey* mark those who refused. **Fig. 2** is also present in a preprint by Hâncean et. al.[24] that explores the assortative mixing in the consumption of processed foods. The preprint is one of the research outputs of the 4P-CAN project, showing that persons who eat processed foods are more likely to have ties to those who exibit similar behaviour. Such results underscore the importance of including network approaches in preventive medicine.

The data collection process took place between 13-30 September 2023. In accordance with ethical norms, prior to each interview, the interviewer prepared a dossier containing the following: *an information sheet* (comprising a brief project introduction, project objectives, organizers and funders of the study, or how personal data would be used), *an informed consent form* (informing the respondent of their voluntary participation in the project and their freedom to withdraw at any time, and confirming their understanding of the information provided in the information sheet), and *a personal data collection and processing form*. We also provided the participants with the contact details (phone numbers and email addresses) of the project director and research team in case they have future queries. Citizens' participation in the project is strictly voluntary and is not based on any monetary system.

The data were collected through face-to-face interviews using the Network Canvas software[27]. Throughout the data collection process, we maintained transparency both in the information provided about the project and in the information entered into the software, as the study participants could see the data that was recorded into the software. Data collection was stopped after ensuring that the individuals from the sample were diverse enough on their socio-demographic profile based on sex, age category (young persons between 18 and 30 years old, adults between 31 and 63, and older adults who were 64 years old or older), education (primary, lower secondary, and tertiary), and personal income (below and above the average national net salary). The number of people in the sample at the end of the data collection process was 83.

The second wave

The second wave of the 4P-CAN PNA study took place between 20-31 March 2024. We interviewed 94 persons, out of which 68 were from the initial panel (built during wave 1, i.e., the pilot phase), resulting in a 72% retention rate. New persons were introduced into the panel to mitigate participant drop-out and attrition. Besides the modifications in the composition of the panel, the differences between the first two waves consist in the focus and themes included in the questionnaire used in the second wave.

The Appendix contains the questionnaire used in the pilot phase (wave 1) (Appendix A1) and the questionnaire developed after (for subsequent waves – 2, 3, 4, etc.) (Appendix A2). The ego-centric questionnaire developed for the pilot phase was extensive, averaging 80 minutes in its applications. After feedback from the citizens, the questionnaire size was reduced. The main differences between the pilot questionnaire and the one used in the second wave are in the number of themes, the method of generating the alters, and the prompt for creating the alter-alter ties. Regarding the themes, the pilot phase used an extended list, including alcohol consumption, exposure to UV radiation, vaccination against COVID-19, HPV, and Hepatitis B, or knowledge about campaigns against cancer risk factors. The second wave questionnaire kept items related to obesity, smoking, physical activity, and had a higher focus on dietary habits, while also introducing new items important for the study of this theme (e.g. , when they eat, how often they.

Regarding alter elicitation methods, the pilot questionnaire used a single name generator: *Please nominate 25 people (18 years old plus) you interact with (or meet). You can start with the people you interact with most often. These may be family members, friends, acquaintances, neighbours, work colleagues, etc.* In the second wave, several name generators were used: 1) *household members*; 2) *other family members and relatives with whom they interact face-to-face at least weekly*; 3) *co-workers with whom they interact face-to-face at least weekly*; and 4) *friends with whom they interact face-to-face at least weekly*. We chose to modify the alter elicitation method based on the fact that social influence happens through strong ties and repeated interaction, and to reduce the risk of capturing incorrect information about the alters.

In terms of alter-alter ties, the pilot questionnaire looked at the emotional closeness between alters (*close friends*; *simple friends*; *acquaintances*). The second wave questionnaire focused on the frequency of interactions between alters (if they *see each other at least once a week* or *less than once a week*). Given

the specificity of the PNA method, the alter-alter ties were captured from the ego's perspective.

External validation

The PNA methodology proposed in the 4P-CAN project was elaborated in collaboration with international experts in social network analysis. On the 8th of December, 2023, the workshop "Network analysis and public health" was held in Bucharest to discuss applications of network science methods in the study of public health and primary prevention. During this workshop, the 4P-CAN research team gathered feedback on the execution of the project's initial phase, received scientific guidance in building the questionnaires for the next waves, and discussed further methodological applications.

A panel of experts employed both face and content validity of questionnaire items, including José Luis Molina and Jürgen Lerner. José Luis Molina is a full professor at the Department of Social and Cultural Anthropology, at the Autonomous University of Barcelona, coordinator of the research group GRAFO - Grup de Recerca en Antropologia Fonamental i Orientada, and also serves as the president of the University's Research Ethics Committee. Since 2007, he has been the organizer of the international winter/summer course "Personal Networks: Theory, Methods and Applications", held at the Autonomous University of Barcelona. His scientific interests span methodological[28], ethical[29,30], and substantial developments of the personal network analysis method, including the impact of the COVID-19 pandemic on vaccination hesitancy[31] and human interactions[32]. Jürgen Lerner is a researcher in computer science at the Department of Computer and Information Science, at the University of Konstanz. His primary research interests include social network analysis, computational social science, and statistical network modeling. Lerner is particularly enthusiastic about network analysis applications across social sciences and interdisciplinary research. As a network scientist, Jürgen Lerner is constantly working on developing the relational hyperevent modeling framework[33,34] (RHEM), a novel way for the analysis of dynamic networks. RHEM has been found applicable for various subjects, from contact diaries[35] to the spread of the SARS-CoV-2 virus following the age[36] and occupation[37] characteristics.

During the workshop, the methodological quality and applicability of the study were assessed. Initially, the panel validated the questionnaire, which focused on gathering information using the PNA method. This validation process was important to confirm that the methodological tools were appropriate and accurate for the research objectives. Secondly, we successfully gathered specialized feedback. Given Professor Molina's foundational research background and Jürgen Lerner's expertise in statistical modeling of network analysis data, their

feedback was essential in continuously improving research tools and methods. Thirdly, considering Professor Molina's role as the Head of the Ethical Committee at the University of Barcelona, we aimed to ensure that our study adhered to high ethical and academic standards. The meeting with the panel of experts was important for confirming that the study is well-grounded, methodologically sound, and relevant in the context of personal network influences on individuals' health behaviors and perceptions. The feedback received was significant for maintaining the relevance and accuracy of the research, particularly as data will be collected in multiple waves.

Figure 3 Project team members together with the international experts from the panel (part I)

Note: from left to right - Simona Puncioiu (research team member), Bianca Cucoș (Center for Innovation in Medicine), Isabela Tincă (research team member), Bogdan Vidrașcu (research team member), Iulian Oană (research team member), Bianca-Elena Mihăilă (research team member), Jürgen Lerner (University of Konstanz), José Luis Molina (Autonomous University of Barcelona), Florin Găină (research team member), Marius Geantă (Head of the Center for Innovation in Medicine), and Marian-Gabriel Hâncean (Professor and Head of the Romanian research team).

Figure 4 Project team members together with the international experts from the panel (part II)

Note: from left to right - Florin Găină (research team member), Bianca-Elena Mihăilă (research team member), Bogdan Vidrașcu (research team member), Simona Puncioiu (research team member), Marian-Gabriel Hâncean (Professor and Head of the Romanian research team), Iulian Oană (research team member), Jürgen Lerner (University of Konstanz), José Luis Molina (Autonomous University of Barcelona), Isabela Tincă (research team member), Bianca Cucoș (Center for Innovation in Medicine), Marius Geantă (Head of the Center for Innovation in Medicine).

References

1. WHO. Cancer. (2022). https://www.who.int/news-room/fact-sheets/detail/cancer. Accessed on 12 October 2023.
2. EC. News announcement. Cancer in Europe: 5 things the data tells us. 12 October 2022. https://joint-research-centre.ec.europa. eu/jrc-news-and-updates/ Accessed on 12 October 2023.
3. INOMED. Raportul de garda. (2023). https://raportuldegarda.ro/romania-profil-cancer-2023-diferenta-rata-mortalitate-media-ue-crestere/. Accessed on 2 October 2023.
4. GoINVO. Determinants of health. (2020). https://www.goinvo.com/vision/determinants- of- health/. Accessed on 12 October 2023.
5. 4P-CAN. 4P-CAN. https://4p-can.eu (2024).
6. Silk, M. J. et al. The application of statistical network models in disease research. Methods Ecol Evol 8, 1026–1041 (2017).
7. Molina, J. L. & Borgatti, S. P. Moral bureaucracies and social network research. Social Networks 67, 13–19 (2021).
8. Christakis, N. A. & Fowler, J. H. The Collective Dynamics of Smoking in a Large Social Network. N Engl J Med 358, 2249–2258 (2008).
9. Christakis, N. A. & Fowler, J. H. The Spread of Obesity in a Large Social Network over 32 Years. N Engl J Med 357, 370–379 (2007).
10. Fernández-Peña, R., Ovalle-Perandones, M.-A., Marqués-Sánchez, P., Ortego-Maté, C. & Serrano-Fuentes, N. The use of social network analysis in social support and care: a systematic scoping review protocol. Syst Rev 11, 9 (2022).
11. Marin, A. & Wellman, B. Social Network Analysis: An Introduction. in The SAGE Handbook of Social Network Analysis (eds. Scott, J. & Carrington, P. J.) 11–25 (SAGE Publications Ltd, London, 2011).
12. Borgatti, S. P., Everett, M. G. & Johnson, J. C. Analyzing Social Networks. (SAGE Publications Ltd, London, 2013).
13. Scott, J. Social Network Analysis: A Handbook. (SAGE Publications Ltd, London, 2000).
14. McCarty, C., Lubbers, M. J., Vacca, R. & Molina, J. L. Conducting Personal Network Research: A Practical Guide. (Guilford Publications, 2019).
15. Ripley, R. M., Snijders, T. A., Boda, Z., Vörös, A. & Preciado, P. Manual for RSIENA. University of Oxford, Department of Statistics, Nuffield College 1, 2011 (2011).

16. Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M. & Morris, M. statnet: Software tools for the representation, visualization, analysis and simulation of network data. Journal of statistical software 24, 1548 (2008).
17. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1–48 (2015).
18. Zhang, S., De La Haye, K., Ji, M. & An, R. Applications of social network analysis to obesity: a systematic review. Obesity Reviews 19, 976–988 (2018).
19. van Woudenberg, T. J. et al. Identifying Influence Agents That Promote Physical Activity Through the Simulation of Social Network Interventions: Agent-Based Modeling Study. J Med Internet Res 21, e12914 (2019).
20. Sawka, K. J., McCormack, G. R., Nettel-Aguirre, A., Hawe, P. & Doyle-Baker, P. K. Friendship networks and physical activity and sedentary behavior among youth: a systematized review. Int J Behav Nutr Phys Act 10, 130 (2013).
21. Schaefer, D. R., Adams, J. & Haas, S. A. Social networks and smoking: exploring the effects of peer influence and smoker popularity through simulations. Health Educ Behav 40, 24S-32S (2013).
22. Blok, D. J., De Vlas, S. J., Van Empelen, P. & Van Lenthe, F. J. The role of smoking in social networks on smoking cessation and relapse among adults: A longitudinal study. Preventive Medicine 99, 105–110 (2017).
23. Keim-Klärner, S. et al. Social inequality, social networks, and health: a scoping review of research on health inequalities from a social network perspective. International Journal for Equity in Health 22, 74 (2023).
24. Marian-Gabriel Hâncean et al. Processed food intake assortativity in the personal networks of East European older adults. medRxiv 2024.01.25.24301787 (2024) doi:10.1101/2024.01.25.24301787.
25. Hâncean, M.-G., Lubbers, M. J. & Molina, J. L. Measuring transnational social fields through binational link-tracing sampling. PLoS One 16, e0253042 (2021).
26. Molina, J. L., Lubbers, M. J., Hâncean, M.-G. & Fradejas-García, I. Short Take: Sampling from Transnational Social Fields. Field Methods 34, 256–264 (2022).
27. Complex Data Collective. Network Canvas: Interviewer. Zenodo https://doi.org/10.5281/ZENODO.6026548 (2023).
28. McCarty, C. & Lubbers, M. J. *Conducting Personal Network Research: A Practical Guide*. (The Guilford Press, New York City, 2019).
29. Borgatti, S. P. & Molina, J.-L. Toward ethical guidelines for network research in organizations. *Social Networks* **27**, 107–117 (2005).
30. Molina, J. L. & Borgatti, S. P. Moral bureaucracies and social network research. *Social Networks* **67**, 13–19 (2021).

31. Hâncean, M.-G., Lerner, J., Perc, M., Molina, J. L. & Geantă, M. Assortative mixing of opinions about COVID-19 vaccination in personal networks. *Sci Rep* **14**, 3385 (2024).
32. Fradejas-García, I., Lubbers, M. J., García-Santesmases, A., Molina, J. L. & Rubio, C. Etnografías de la pandemia por coronavirus: emergencia empírica y resignificación social. *Perifèria. Revista d'investigació i formació en Antropologia* **25**, 4–21 (2020).
33. Lerner, J., Tranmer, M., Mowbray, J. & Hancean, M.-G. REM beyond dyads: relational hyperevent models for multi-actor interaction networks. *arXiv preprint arXiv:1912.07403* (2019).
34. Lerner, J. & Lomi, A. A dynamic model for the mutual constitution of individuals and events. *Journal of Complex Networks* **10**, cnac004 (2022).
35. Lerner, J., Lomi, A., Mowbray, J., Rollings, N. & Tranmer, M. Dynamic network analysis of contact diaries. *Social Networks* **66**, 224–236 (2021).
36. Hâncean, M.-G. *et al.* The role of age in the spreading of COVID-19 across a social network in Bucharest. *J Complex Netw* **9**, (2021).
37. Hâncean, M.-G. *et al.* Occupations and their impact on the spreading of COVID-19 in urban communities. *Sci Rep* **12**, 14115 (2022).
38. World Health Organization. *WHO STEPS surveillance manual: the WHO STEPwise approach to chronic disease risk factor surveillance* (2017). https://cdn.who.int/media/docs/default-source/ncds/ncd-surveillance/steps/stepsmanual.pdf?sfvrsn=c281673d_8. Accessed on 17 June 2023.
39. Harris, C. V., Bradlyn, A. S., Coffman, J., Gunel, E. & Cottrell, L. BMI-based body size guides for women and men: development and validation of a novel pictorial method to assess weight-related concepts. Int J Obes 32, 336–342 (2008).

Appendix

A1. Ego-centric questionnaire used in the pilot phase (September 2023)

I. Data about ego (the study participant)

I.1. Sociodemographic data

[Interviewer: The following text is a speech template you can use at the beginning of the interview to provide the respondent with general information about the questionnaire's duration and structure. Depending on the context, you can adapt the key ideas in a different form or omit them if not necessary.
*We will start with some general questions, such as your date of birth, education, occupation, whether you are married, and so on. Following that, we will proceed to ask more specific questions, including your medical history, weight, height, smoking habits, dietary preferences, and so forth. After that, we will continue with questions about persons in your life for whom we will request similar information. Completing this questionnaire may take an hour and a half to two hours. We can take a break whenever you wish and then return to the questions***.]**

#1 **sd_ename** What is your full name?

_____
[Open-ended question. Text format.
Note: *First, write down the **last name(s)**, then the **first name(s)**. Request the **full name**. Ask for clarification if it is not clear*.**]**

#2 **sd_ephone** Please provide me with a phone number so that we can stay in contact during the study:

_____
[Open-ended question. Numeric format.
Note: *Request the country code if necessary. Do not write anything if they do not want to respond*.**]**

#3 **sd_edob** What is your date of birth?

_____
[Open-ended question. Text format.
Note: *Write the date of birth using the format **DD-MM-YYYY***. Do not write anything if they do not want to respond.**]**

#4 **sd_esex** Respondent's sex:
_1. Male
_2. Female
[Close-ended question. Binary format.
Note: *Do not read. Do automatic identification*.**]**

#5 **sd_eedu** What is the highest grade or level of school you graduated?

_1. Did not attend school
_2. Incomplete primary education
_3. Primary education
_4. Incomplete lower secondary education
_5. Lower secondary education
_6. Arts & Crafts school (upper secondary education – vocational)*
_7. Ten obligatory years (upper secondary education – partial level completion)*
_8. Highschool – unfinished (upper secondary education – partial level completion)
_9. Highschool – finished (with diploma)
_10. Post-secondary or trade school (non-tertiary education)
_11. Bachelor's degree or equivalent level
_12. Master's degree or equivalent level
_13. Ph.D. or equivalent level
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Ordinal format.
Note: *Allow the respondent to answer freely and categorize the response accordingly.* ***Read the categories only if necessary.**** **Specific categories in the Romanian educational system.]**

#6 sd_eocp What is your current occupation? / What was the last occupation you had?

_____
[Open-ended question. Text format.
Note: *Write down the respondent's occupation. If they are not employed, note whether they are a homemaker, a daily-worker, a seasonal labourer, etc. If retired, ask if it's due to illness, and if so, ask what illness. If unemployed, ask if they receive unemployment benefits.* ***Request specific, detailed answers. If the person provides vague responses, seek clarification****. For example, if they say 'teacher,' ask what kind of teacher – elementary school, middle school, etc. Do not write anything if they do not want to respond.*]

#7 sd_ems At the moment, you are…?:
_1. Single
_2. Married
_3. Unmarried, but in a stable relationship
_4. Another situation. Please detail:

_____
[-99.] DK / NA [Interviewer: Do not read]
[Close-ended question. Nominal format with a **single** response option.
Note: *Read the response options and categorize the respondent accordingly. Adapt the response options to respondent's sex (if the language requires it).*]

#8 sd_ehp In total, how many people live with you (spouse, children, parents, other relatives, friends, etc.)?

_____
[Open-ended question. Numeric format.
Note: *Help the respondent calculate the total if needed.* ***Write down the total number of people, including the respondent****. Don't write anything down if they don't want to answer.*]

#9 sd_epi Currently, your monthly personal income is …?

_1. Under 500 RON (≈ 100 EUR; 1 EUR ≈ 5 RON)
_2. Between 500 and 1000 RON (100-200 EUR)
_3. Between 1001 and 1500 RON (201-300 EUR)
_4. Between 1501 and 2000 RON (301-400 EUR)
_5. Between 2001 and 2500 RON (401-500 EUR)
_6. Between 2501 and 3000 RON (501-600 EUR)
_7. Between 3001 and 3500 RON (601-700 EUR)
_8. Between 3501 and 4000 RON (701-800 EUR)
_9. Between 4001 and 5000 RON (801-1000 EUR)
_10. Between 5001 and 6000 RON (1001-1200 EUR)
_11. Between 6001 and 7000 RON (1201-1400 EUR)
_12. Between 7001 and 8000 RON (1401-1600 EUR)
_13. Between 8001 and 9000 RON (1601-1800 EUR)
_14. Between 9001 and 10000 RON (1801-2000 EUR)
_15. Over 10000 RON (Over 2000 EUR)
[-99.] DK / NA [Interviewer: Do not read]
[Close-ended question. Ordinal format.
Note: ***Don't read the response options.*** *Let the respondent provide an open answer and place it in the appropriate category. Read answer options 1 to 15 only if the respondent is hesitant to provide a specific answer. Convert to common (popular) monetary expressions if needed.*]

#10 **sd_ehi** Currently, the monthly income in your household is …?
_1. *Under 500 RON (≈ 100 EUR; 1 EUR ≈ 5 RON)*
_2. Between 500 and 1000 RON (100-200 EUR)
_3. Between 1001 and 1500 RON (201-300 EUR)
_4. Between 1501 and 2000 RON (301-400 EUR)
_5. Between 2001 and 2500 RON (401-500 EUR)
_6. Between 2501 and 3000 RON (501-600 EUR)
_7. Between 3001 and 3500 RON (601-700 EUR)
_8. Between 3501 and 4000 RON (701-800 EUR)
_9. Between 4001 and 5000 RON (801-1000 EUR)
_10. Between 5001 and 6000 RON (1001-1200 EUR)
_11. Between 6001 and 7000 RON (1201-1400 EUR)
_12. Between 7001 and 8000 RON (1401-1600 EUR)
_13. Between 8001 and 9000 RON (1601-1800 EUR)
_14. Between 9001 and 10000 RON (1801-2000 EUR)
_15. Over 10000 RON (Over 2000 EUR)
[-99.] DK / NA [Interviewer: Do not read]
[Close-ended question. Ordinal format.
Note: ***Don't read the response options.*** *Let the respondent provide an open answer and place it in the appropriate category. Read answer options 1 to 15 only if the respondent is hesitant to provide a specific answer.* ***If the ego lives alone, check the answer given for the personal income****. Convert to common (popular) monetary expressions if needed.*]

I.2. Medical data

#11 mr_ecp Are you suffering or have you suffered from any cancer type? [⇒ *Go to question 13 if the respondent says that they don't or haven't suffered from cancer*]

_1. No

_2. Yes, I currently suffer from cancer

_3. I am a cancer survivor

[-99.] DK / NA [Interviewer: do not read]

[Close-ended question. Nominal format with a single response option.]

#12 mr_ecpt What type of cancer are you suffering from / have you suffered from?

_____

[Open-ended question. Text format.

Note: *Do not write anything if they don't want to answer.***]**

#13 mr_esd Are you currently suffering from any illness? (Explicitly ask for the type of illness/illnesses, if applicable)

_____

[Open-ended question. Text format.

Note: *Write down "NO" if they don't suffer from any illness or if they state that they don't know. Do not write anything if they don't want to answer.*]

#14 mr_eldvd When was the last time you visited a doctor for yourself? **[**⇒ *Go to question 18 if the respondent says they never went to the doctor***]**

_____

[Open-ended question. Text format.

Note: *Write down "NO" or "NEVER" if the respondent says they never went to the doctor.* *Do not write anything if they don't want to answer.***]**

#15 mr_eldvl In what locality did you go when you went (last time) to the doctor?

_____

[Open-ended question. Text format.

Note: *Write down the location as precisely as possible.* *Do not write anything if they don't want to answer or never went to the doctor.***]**

#16 mr_eldvhd When did you last go to the doctor for a heart check-up?

_____

[Open-ended question. Text format.

Note: *Write down "NO" or "NEVER" if the respondent says they never went to the doctor for a heart check-up. Do not write anything if they don't want to answer.*]

#17 mr_eldvhl In what locality you went (last time) for the heart check-up?

_____

[Open-ended question. Text format.

Note: *Do not write anything if they don't want to answer or never went to the doctor for a heart check-up.*]

#18 **mr_efcan** Is there anyone in your family who suffered or currently suffers from cancer? If so, who and what type of cancer?

_____

[Open-ended question. Text format.
Note: *If 'Yes', write down **the kinship relation and the cancer type**. Write down all the answers provided by the respondent. Example: mother – breast cancer; father – lung cancer. Do not write anything if they don't want to answer.*]

#19 **mhs_efcvd** Is there anyone in your family who suffered or currently suffers from heart disease? If so, who and what kind of disease?

_____

[Open-ended question. Text format.
Note: *If 'Yes', write down the **kinship relation and disease type**. Write down all the answers given by the respondent. Do not write anything if they don't want to answer.*]

#20 **mr_efdn** What is the name of your family doctor?

_____

[Open-ended question. Text format.
Note: *Do not write anything if they don't want to answer.*]

#21 **mr_ev_cov1** Have you been vaccinated against COVID-19?
_1. Yes
_2. No
[-99.] DK/ NA [Interviewer: do not read]
[Close-ended question. Binary format.]

#22 **mr_ev_cov2** Who advised you to get vaccinated against COVID-19?
_1. A family member
_2. A close friend
_3. A work colleague
_4. The family doctor
_5. A doctor, another than the family doctor
_6. Someone from the Hall (local administration)
_7. Professor / Teacher / kindergarten teacher / school's medical personnel
_8. An acquaintance
_9. Public figures from TV or radio
_10. People or pages I follow on the internet
[-98.] Not the case (not vaccinated) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Nominal format with **multiple** response options.]

#23 **mr_ev_hpv1** Have you been vaccinated against HPV? [*Ask the men as well*]
_1. Yes
_2. No
[-99.] DK/ NA [Interviewer: do not read]

[Close-ended question. Binary format.]

#24 **mr_ev_hpv2** Who advised you to get vaccinated against HPV? [*Ask the men as well*]
_1. A family member
_2. A close friend
_3. A work colleague
_4. The family doctor
_5. A doctor, other than the family doctor
_6. Someone from the Hall (local administration)
_7. Professor / Teacher / kindergarten teacher / school's medical personnel
_8. An acquaintance
_9. Public figures from TV or radio
_10. People or pages I follow on the internet
[-98.] Not the case (not vaccinated) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Nominal format with **multiple** response options.]

#25 **mr_ev_hpb1** Have you been vaccinated against Hepatitis B?
_1. Yes
_2. No
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Binary format.]
#26 **mr_ev_hpb2** Who advised you to get vaccinated against Hepatitis B?
_1. A family member
_2. A close friend
_3. A work colleague
_4. The family doctor
_5. A doctor, another than the family doctor
_6. Someone from the Hall (local administration)
_7. Professor / Teacher / kindergarten teacher / school's medical personnel
_8. An acquaintance
_9. Public figures from TV or radio
_10. People or pages I follow on the internet
[-98.] Not the case (not vaccinated) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Nominal format with **multiple** response options.]

#27 **mr_einfo** What are your sources of information about health and prevention (lifestyle) in particular?
_1. Central TV stations
_2. Local press
_3. Online, using Google, Bing, or other search engines
_4. Online, using social media - Facebook, TikTok, Instagram etc.
_5. Online influencers
[-99.] DK / NA [Interviewer: do not read]

[Close-ended question. Nominal format with **multiple** response options.]

I.3. Body-Mass Index

#28 **bmi_eow** Person's weight according to the scale:

_____

[Open-ended question. Text format.
Note: *Write down the weight indicated by the scale using two decimals – e.g., 78.29. Do not write anything if they don't want or cannot be weighted.*]

#28bis1 **bmi_esw** What is your weight? [If they don't want or cannot be weighted]

_____

[Open-ended question. Text format.
Note: *Write down the answer in kilograms and grams. – e.g., 60 (if their answer is 60 kilograms); 70.5 (if their answer is 70 kg and a half). Do not write anything if they don't want to answer.*]

#28bis2 **bmi_eswp** What was your weight before getting pregnant? [*If the respondent is pregnant*]

_____

[Open-ended question. Text format.
Note: *Write down the answer in kilograms and grams. – e.g., 60 (if their answer is 60 kilograms); 70.5 (if their answer is 70 kg and a half). Do not write anything if the respondent is not pregnant or if they don't want to answer.*]

#29 **bmi_eoh** Person's height according to measurement:

_____

[Open-ended question. Text format.
Note: *Write down the height indicated by the measuring instrument using two decimals – e.g., 1.75. Do not write anything if they don't want or cannot have their height measured.*]

#29bis1 **bmi_esh** How tall are you? [If they don't want or cannot be measured]

_____

[Open-ended question. Text format.
Note: *Write down the height given by the respondent using two decimals – e.g., 1.75. Do not write anything if they don't want to answer.*]

I.4. Smoking

#30 **smk_ess** Currently, you are:
_1. A smoker
_2. An occasional smoker (at events)
_3. A former smoker
_4. I've smoked very few cigarettes to consider myself a smoker [⇒ *Go to question 34*]
_5. I've never smoked [⇒ *Go to question 34*]
[-99.] DK / NA [Interviewer: do not read] [⇒ *Go to question 34*]

[Close-ended question. Nominal format with a **single** response option.]

 # 31 smk_est What kind of cigarettes do you / did you smoke? [*Current and former smokers; Q30 = 1, 2, or 3*]
_1. Filter cigarettes
_2. Cigarettes without filter
_3. Rolling tobacco
_4. Electronic cigarettes (heated tobacco)
_5. Vaping devices (nicotine liquid devices)
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Nominal format with **multiple** response options.]

#32 smk_esage At what age did you start smoking? [*Only current smokers; Q30 = 1 or 2*]

_____
[Open-ended question. Text format.
Note: *Don't write anything if it's not the case (non-smoker) or they don't want to answer.*]

#32bis smk_esper How many years have you smoked? [*Only former smokers; Q30 = 3*]

_____
[Open-ended question. Text format.
Note: *Don't write anything if it's not the case (non-smoker) or they don't want to answer.*]

#33 smk_epw Roughly, how many cigarette packs do you smoke / have you smoked on a weekly basis? [*Current and former smokers; Q30= 1, 2, or 3*]

_____
[Open-ended question. Text format.
Note: *note the number of packs as mentioned (e.g., 4 packs, 3 and a half packs). A pack has 20 cigarettes. If they mention a daily number of cigarettes, convert it to a packs per week format (e.g., **'10 cigarettes per day'** becomes **'3 and a half packs'**). Don't write anything down if it's not the case or they don't want to answer.*]

I.5. Alcohol

[Interviewer: Read the next section and show the respondent the showcard for types of drinks.
Next, we will discuss the consumption of various drinks containing alcohol. I'm going to show you pictures of several types of drinks such as beer, wine, champagne, liquor, and other various spirits to have a reference for the following question.]

[Use showcard for types of drinks]

#34 alc_elw In the past week, how many of the following drinks have you consumed (in total)?

_____
[Open-ended question. Text format.
Note: *Show the images related to different types of drinks. Help the respondent calculate the total if needed. Write 0 if they haven't consumed any. Do not write anything if they choose not to answer.*]

I.6. Dietary habits

#35 die_eas How often do you add salt or a salty sauce such as soy sauce to your food right before you eat it or while eating it?[1]

_1. Never
_2. Rarely
_3. Often
_4. Always
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Ordinal format.]

#36 die_epf How frequently do you consume foods similar to the ones depicted in these images?? (in total, not from each)

[Use showcard with processed foods]

_1. Never
_2. Less than once a year
_3. Once a year
_4. A few times a year
_5. Once a month
_6. Every two weeks
_7. Weekly
_8. Daily
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Ordinal format.
Note: *Show card with processed foods.* ***Also mention that the question does not refer only to the foods in the images, but also to similar products. If they have problems responding, ask them what do they eat most often and how frequently.***]

#37 die_epfm What are the reasons you eat foods like the ones I've shown you? [*For those who didn't say "NEVER" at Q36*]
_1. They are affordable
_2. I like the taste
_3. It's easy for me to buy them
_4. I don't have anything else to buy from the stores I go to
_5. To cook less
_6. Because I don't cook
[-98.] Not the case (they don't consume such foods) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Nominal format with **multiple** response options.]

[1] This question was adapted from The WHO STEPS Surveillance Manual[38], p. 528.

#38 die_ecds In a typical week, on how many days of the week do you consume sweet juices or eat sweets? Please do not consider freshly squeezed fruit juices.

_____

[Open-ended question. Numeric format.

Note: *Write down "NO" or "NEVER" if they don't consume sugar-based drinks or sweets*. *Do not write anything if they don't want to answer.*]

#39 die_emt Of the following types of meat, which ones do you consume?

_1. Chicken
_2. Beef
_3. Pork
_4. Fish
_5. Lamb/sheep
_6. Venison
_7. Minced meat
_8. Other
[-98.] Not the case (they don't consume such foods) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]

[Close-ended question. Nominal format with **multiple** response options.]

#40 die_emcp Which of the following meat preparation methods do you prefer?

_1. Grilled
_2. Fried
_3. Oven cooked
_4. Boiled
_5. Smoke preserved
_6. Salt preserved
_7. Lard preserved
_8. Other. Please detail:

_____

[-98.] Not the case (they don't consume such foods) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]

[Close-ended question. Nominal format with **multiple** response options.]

 # 41 die_ems In general, the meat you consume comes from …?

_1. Your own household
_2. The household of relatives/friends
_3. Mini-markets (small, local stores)
_4. Super-markets (Carrefour, Auchan, Kaufland, etc.)
_5. Market
_6. Other. Please detail: _____
[-98.] Not the case (they don't consume such foods) [Interviewer: do not read]
[-99.] DK / NA [Interviewer: do not read]

[Close-ended question. Nominal format with **multiple** response options.]

#42 **die_ewc** In general, where does the water you drink come from?
_1. Well (fountain)
_2. Springs
_3. Public water system (from the sink)
_4. Bottled water (from shops)
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Nominal format with **multiple** response options.]

I.7. Physical activity

[Interviewer: Read the following section and present to the respondent the showcards for physical labor. After presenting them, make sure the respondent understands the difference between intense and moderate physical activities.
*The next section is about movement and physical activity. We will consider both the effort you put into work-related activities and other moments. To make it easier for you to respond, I will present several images depicting 'intense physical activities' and 'moderate physical activities' to provide a reference.***]**
[Use showcards describing types of physical effort]

#43 **pac_ewa** At work, which of the following best describes what you do on a typical day?[2]
[Only for those who work (formally or informally)]
_1. Not performing any working tasks
_2. Mostly sitting or standing
_3. Mostly walking or tasks of moderate physical effort
_4. Mostly heavy labour or physically demanding work
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Ordinal format.
Interviewer: Use showcards with moderate physical effort and physically demanding tasks.]

#44 **pac_eha** At home, which of the following best describes what you do on a typical day?[3]
_1. Not performing any working tasks
_2. Mostly sitting or standing
_3. Mostly walking or tasks of moderate physical effort
_4. Mostly heavy labour or physically demanding work
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Ordinal format.

[2] This question was adapted from the European Health Interview Survey (EHIS Wave 3) Methodological manual (2020), re-edition. Eurostat, *European Health Interview Survey (EHIS Wave 3) Methodological Manual*, (2020), p.87.

[3] *Idem.*

Interviewer: Use showcards with tasks of moderate physical effort and physically demanding.**]**

#45 pac_esh Do you have any sport as a hobby (passion) that you practice? (soccer, running, swimming, etc.). If so, which one? [Walking and/or cycling are also considered]

_____

[Open-ended question. Text format.
Interviewer: *Write down "NO" if they don't practice any type of physical activity*. *Do not write anything if they don't want to answer.***]**

I.8. Exposure to UV radiation

#46 uv_epc When you leave the house, do you use sunscreen (or sunscreen products):
_1. Never
_2. In the summer, when I go to the beach / pool
_3. In the summer, only when it is very hot
_4. In the summer, regardless of the temperature
_5. All the time, regardless of the season
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Ordinal format.]

II. Data about alters (ego's social contacts)

II.1. Generating alters

[**Interviewer**: Read the following:
The following sections pertain to the people you regularly talk to or meet. We will begin by naming these persons, and then we will continue with a series of questions about them. Such information will help us better understand your life context.]

#47 name First, please tell me the names of 25 people you regularly talk to or meet, and who are at least 18 years old. These can be family members, friends, acquaintances, neighbours, work colleagues, etc. Please provide their full names.

_____
[Open-ended question. Text format.
Note: *Record last name(s), followed by first name(s). Where it's not clear, seek clarification.*]

II.2. Alters' status

#48 rels_a This person is…?
[Immediately after the alter is generated]
_1. A close friend
_2. A casual friend
_3. An acquittance
_4. A family member

_5. A partner (for unmarried persons)
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Nominal format with a single response option.
Note: *Ask about each alter (when generated) and place it in the appropriate category.*]

#49 rels_akin 3. What is the kinship relation with this person?
[*Immediately after the alter is generated****]***

_____
[Open-ended question per alter. Text format.
Interviewer: *Don't write anything if the alter it's not kin or if they don't want to answer. For this question, kinship relation also includes wife or husband.*]

II.3. Sociodemographic data

#50 sd_asex Alter's sex:
[*Immediately after the alter is generated****]***
_1. Male
_2. Female
[Close-ended question per alter. Binary format.
Interviewer: *Do not read.* ***Automatically classify based on the first name(s)****. Ask only if the name is ambiguous.*]

#51 sd_aage Please tell me the (approximate) age, in years, for the people you mentioned earlier:
[*Alter enumeration****]***

_____
[Open-ended question per alter. Numeric format.
Interviewer: *Ask about each alter and write the age. Do not write anything if they don't want to answer.*]

#52 sd_aedu Please tell me what is the last education level the mentioned persons graduated from:
[*Alter enumeration****]***
_1. Did not attend school
_2. Incomplete primary education
_3. Primary education
_4. Incomplete lower secondary education
_5. Lower secondary education
_6. Arts & Crafts school (upper secondary education – vocational)*
_7. Ten obligatory years (upper secondary education – partial level completion)*
_8. Highschool – unfinished (upper secondary education – partial level completion)
_9. Highschool – finished (with diploma)
_10. Post-secondary or trade school (non-tertiary education)
_11. Bachelor's degree or equivalent level
_12. Master's degree or equivalent level
_13. Ph.D. or equivalent level
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Ordinal format.

Interviewer: *Ask about each alter and place them in the appropriate category.*
*** Romanian specific education category.]**

#53 sd_ams For the mentioned persons, please tell me their marital status:
[*Alter enumeration****]***
_1. Single
_2. Married
_3. Unmarried, but in a stable relation
_4. Another situation. Please detail:

_____
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Nominal format with a **single** response option.
Interviewer: *Ask about each alter and place them in the appropriate category.*]

#54 sd_ashhe Does this person live with you?
[Alter enumeration]
_1. Yes [they live in the same household with ego]
_2. No [they don't live in the same household with ego]
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Binary format.]

II.4. Body-Mass Index

#55 bmi_abmic Please, take a look at the following showcards displaying different body types for men and women. Among the categories you see, please associate the individuals mentioned earlier with a body type. **[***Use showcard which show only letters, not the BMI category***]**
[Alter enumeration]
[**Use showcards with BMI types – women and men**]

_1. A
_2. B
_3. C
_4. D
_5. E
_6. F
_7. G
_8. H
_9. I
_10. J
[-99.] DK / NA [Interviewer: Do not read]
[Close-ended question per alter. Ordinal format.
Note: *Ask about each alter and place them in the appropriate category.*]

II.5. Smoking

#56 smk_ssa Of the persons you mentioned, please tell me which of them are:
[Alter enumeration]
_1. A smoker
_2. An occasional smoker (at events)
_3. A former smoker
_4. They've smoked very few cigarettes to consider them a smoker
_5. They've never smoked
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Nominal format with a single response option.
Note: *Ask about each alter and place them in the appropriate category.*]

II.6. Alcohol

#57 alc_aac Would you say about [ALTER NAME] that:
[Alter enumeration]
_1. They never drink
_2. They don't drink anymore
_3. They drink rarely
_4. They drink often
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Ordinal format.
Note: *Ask about each alter and place them in the appropriate category.*]

II.7. Dietary habits

#58 die_akd Of the people you mentioned, about whom do you know is on a diet?
[Alter enumeration]
_1. Yes [they are on a diet]
_2. No [they are not on a diet]
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Binary format.
Note: *Ask about each alter and place them in the appropriate category.*]

#59 die_apf For the people you've mentioned, I would like you to tell me how often they eat foods like the ones in the following images?
[Alter enumeration]
[**Use showcard with processed foods**]

_1. Never
_2. Less than once a year
_3. Once a year
_4. A few times a year
_5. Once a month

_6. *Every two weeks*
_7. *Weekly*
_8. *Daily*
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Ordinal format.
Note: *Ask about each alter and place them in the appropriate category.*]

II.8. Physical activity

#60 **pac_apab** Of the persons you mentioned, who has a sport as hobby (passion) which they practice? If so, what sport or activity? [Walking and/or cycling are also considered]
[Alter enumeration]

_____
[Open-ended question. Text format.
Note: *Write down "NO" if they don't practice any type of physical activity. Do not write anything if they don't want to answer.*]

II.9. Medical data

#61 **mr_avac_all** Has this person been vaccinated against…?
[*Alter enumeration****]***

1.	*COVID-19*	*Yes*
-1.	*COVID-19*	*No*
-11.	*COVID-19*	*DK/NA*
2.	*HPV*	*Yes*
-2.	*HPV*	*No*
-22.	*HPV*	*DK/NA*
3.	*Hepatitis B*	*Yes*
-3.	*Hepatitis B*	*No*
-33.	*Hepatitis B*	*DK/NA*

[Close-ended question per alter.
Note: *Due to software limitations, the question has a nominal format with* ***multiple*** *response options, but a* ***single*** *option per vaccine type.*]

II. 10. Ego-Alter ties

#62 **ego_alter_meet** How often do you typically meet with the people you previously mentioned?
[*Alter enumeration****]***
_1. *Less than once a year*
_2. *Once a year*
_3. *A few times a year*
_4. *Once a month*
_5. *Every two weeks*

_6. Weekly
_7. Daily
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Ordinal format.
Note: *Ask about each alter and place them in the appropriate category.*]

#63 **ego_alter_emo** How emotionally close do you feel to the people you previously mentioned?
[Alter enumeration]
_1. Not at all close
_2. Not very close
_3. Close
_4. Very close
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question per alter. Ordinal format.
Note: *Ask about each alter and place them in the appropriate category.*]

II.11. Alter-Alter ties

#64 **aatie_know** Please tell me if these persons, even if they are related to each other, are:
[Dyad enumeration]

_0. They don't know each other / DK-NA [Interviewer: do not read]
_1. Acquittances
_2. Casual friends
_3. Close friends
[Close-ended question per dyad. Ordinal format.]

III. Ego – supplementary data

III.1. Information about children

#65 **sd_ech** How many children do you have? [Proceed to question 69 if they don't have children]

_____
[Open-ended question. Numeric format.
Note: *Complete with numerical answers. Write 0 if the person has no children. Write 0 including if the person is a pregnant woman without other children. Don't write anything down if they don't want to answer.*]

#66 **sd_echm** Of your children, how many are under 18?

_____

[Open-ended question. Numeric format.
Note: *Adjust the question's wording if the respondent has only one child. Write 0 if the person has no children under 18. Don't write anything down if they don't want to answer.*]

#67 **sd_echh** Of your children, how many live with you?

_____
[Open-ended question. Numeric format.
Note: *Adjust the question's wording if the respondent has only one child. The question refers to all children, not just minors. Write 0 if the person has no children living with them in the same household. Don't write anything down if they don't want to answer.*]

#68 **sd_echv** Of your children, which were vaccinated against COVID-19, HPV, and Hepatitis B?

_____
[Open-ended question. Text format.
Note: *Adjust the question's wording if the respondent has only one child. The question refers to all children, not just minors. Write down the sex, age, and type of vaccine(s) (COVID-19, HPV, HBV). Do not note anything if the person chooses not to answer.*]

III.2. Campaigns – risk factors

{Smoking}

#69 **smk_koc1** In the last five years, do you remember any prevention campaign regarding smoking? *[Campaigns include anything: actions, TV messages, label messages, etc.]*
_1. Yes
_2. No
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Binary format.]

#70 **smk_koc2** Who organized this / these campaign(s)? [*Only for those who know*]

_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, don't remember, or don't want to answer.*]

#71 **smk_koc3** In your opinion, who should organize such campaigns?

_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, or don't want to answer.*]

{Alcohol}

#72 **alc_koc1** In the last five years, do you remember any prevention campaigns regarding excessive alcohol consumption? [*Campaigns include anything: actions, TV messages, label messages, etc.*]
_1. Yes
_2. No
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Binary format.]

#73 **alc_koc2** Who organized this / these campaign(s)? [Only for those who know]
_____
[Open-ended question. Text format.
Interviewer: *Don't write anything if they don't know, don't remember, or don't want to answer.*]

#74 **alc_koc3** In your opinion, who should organize such campaigns?
_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, or don't want to answer.*]

{Dietary habits}

#75 **die_koc1** In the last five years, do you remember any campaigns related to healthy eating? *[Campaigns include anything: actions, TV messages, label messages, etc.]*
_1. Yes
_2. No
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Binary format.]

#76 **die_koc2** Who organized this / these campaign(s)? [Only for those who know]
_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, don't remember, or don't want to answer.*]

#77 **die_koc3** In your opinion, who should organize such campaigns?
_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, or don't want to answer.*]

{Physical activity}

#78 **pac_koc1** In the last five years, do you remember any campaigns regarding the benefits of sport or physical activity? *[Campaigns include anything: actions, TV messages, label messages, etc.]*

_1. Yes
_2. No
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Binary format.]

#79 **pac_koc2** Who organized this / these campaign(s)? [Only for those who know]

_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, don't remember, or don't want to answer.*]

#80 **pac_koc3** In your opinion, who should organize such campaigns?

_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, or don't want to answer.*]

{UV radiations}

#81 **uv_koc1** In the last five years, do you remember any campaigns regarding the protection against sun? *[Campaigns include anything: actions, TV messages, label messages, etc.]*
_1. Yes
_2. No
[-99.] DK / NA [Interviewer: do not read]
[Close-ended question. Binary format.]

#82 **uv_koc2** Who organized this / these campaign(s)? [Only for those who know]

_____
[Open-ended question. Text format.
Interviewer: *Don't write anything if they don't know, don't remember, or don't want to answer.*]

#83 **uv_koc3** In your opinion, who should organize such campaigns?

_____
[Open-ended question. Text format.
Note: *Don't write anything if they don't know, or don't want to answer.*]

{Cancer prevention campaigns}

84 **info_ecs** From which organizations or institutions would you prefer to receive information regarding cancer prevention?
_1. Central authorities (e.g. Ministry of Health, national authorities for health)
_2. Local authorities (local level institutions – Hall, local authorities for health)
_3. Doctors (hospital doctors, family doctors, etc.)
_4. Medical assistants / Nurses
_5. Press organizations
_6. NGOs such as foundations and associations involved with this subject
_7. Other organisations (schools, churches, etc.)
[-99.] DK / NA [Interviewer: do not read]

[Close-ended question. Nominal format with **multiple** response options.]

IV. End of interview data

Can you help us contact at least five people (your family members, friends or acquaintances) who are at least 18 years old and currently living in [**PLACE OF LIVING-LAB**], to participate in our study? For that, we would need their name and phone number. Kindly initiate contact with them in advance.

	Last name and first name(s) (**#85 name**)	Phone number (**#86 con_phone**)
REF 1		
REF 2		
REF 3		
REF 4		
REF 5		

#87 fin_obs Other observations:

_____

[Open-ended question. Text format.

Note: *Write down observations that can be important (e.g., what the respondent says about calling them to see if they talked with their contacts). Do not write anything if it's not the case.*]

Showcards used in the pilot phase

Figure 5 Showcard with different types of food

Figure 6 Showcard with different types of spirits

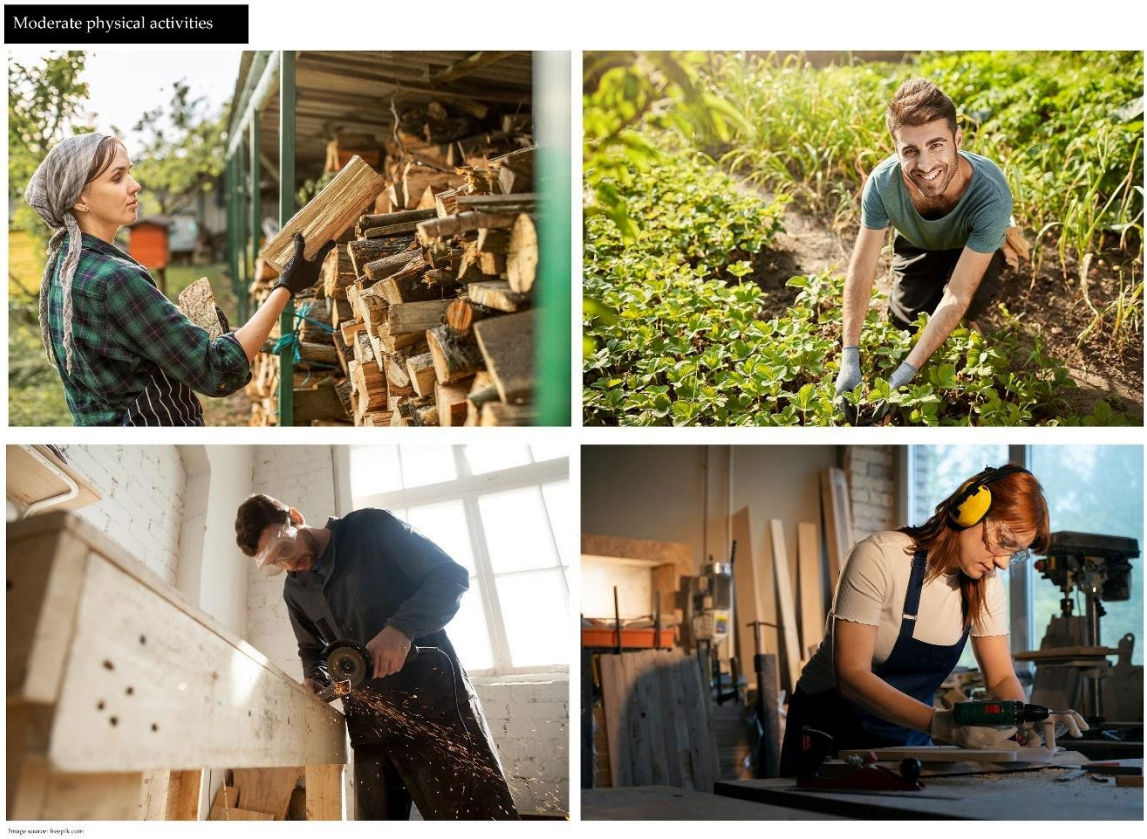

Figure 7 Showcard with moderate physical activities

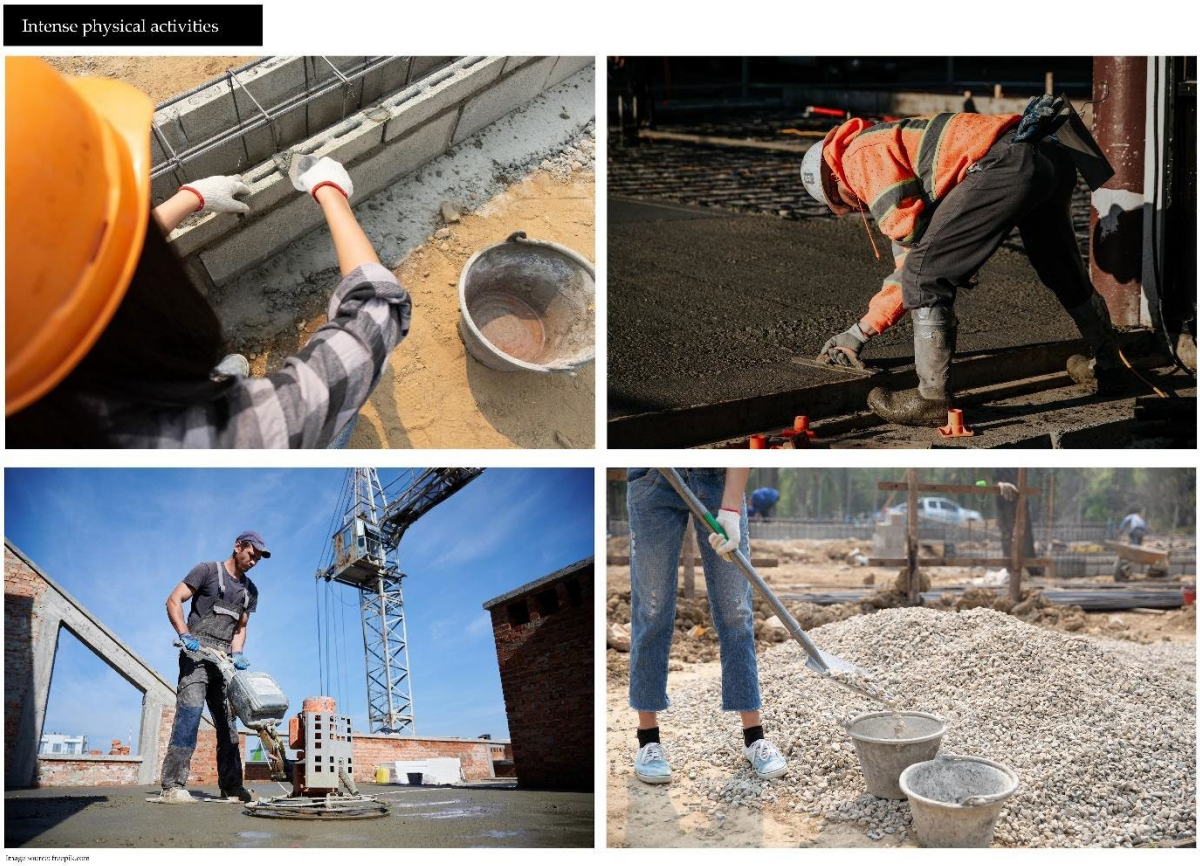

Figure 8 Showcard with intense physical activities

Figure 9 Showcard with different types of bodies for females
Note: The showcard is extracted from the work of Harris et al.[39]

Weight by Measured BMI	Under weight	Normal weight	Over weight	Class I Obesity	Class II Obesity	Class III Obesity

Figure 10 Showcard with different types of bodies for females grouped by BMI categories
Note: The showcard is extracted from the work of Harris et al.[39]

Figure 11 Showcard with different types of bodies for males
Note: The showcard is extracted from the work of Harris et al.[39]

Weight by Measured BMI	Under weight	Normal weight	Over weight	Class I Obesity	Class II Obesity	Class III Obesity

Figure 12 Showcard with different types of bodies for males grouped by BMI categories.
Note: The showcard is extracted from the work of Harris et al.[39]

A2. Ego-centric questionnaire used in wave two (March 2024)

I. Data about ego (the study participant)

I.0 Respondent type

#0 status.respondent Place the person within a category:
_a. new participant
_b. panel participant
[Closed response. Binary format.
Note: *participants from the previous wave (_b_) will not respond to questions marked with _N_.*]

I.1 Socio-demographic data

#1 ego.name What is your full name?

_____
[Open response. Text format.
Note: *First, write the last name, then the first name. Ask for the full name AS STATED IN THE ID CARD. Seek clarifications if unclear.*]

#2 ego.phone Please provide me with a phone number so that we can stay in contact during the study:

_____
[Open response. Numeric format.
Note: *For respondents who participated in the previous wave, confirm the phone number. For new participants, record it in the format +COUNTRY PREFIX XXXXXXXXX.*]

#3N ego.birth When were you born (day, month, and year)?

_____
[Open response. Text format.
Note: *Write the date of birth in the form DD-MM-YYYY. Do not record anything if they do not wish to respond. This question applies only to new participants.*]

#4N ego.sex Participant's sex:
_1. Male
_2. Female
[Closed response. Binary format.
Note: *Do not read. Automatically categorize the participant's sex. This question applies only to new participants.*]

#5N ego.edu What is the highest level of education you have completed?

_1. Up to eight years of education [finished lower secondary education]
_2. Between nine years and post-secondary school/trade school [non-tertiary education]
_3. University studies (minimum bachelor's degree)
[-99.] Don't know / No response [Note: do not read]
[Closed response. Ordinal format.
Note: *Let the respondent provide an open answer and place it in the appropriate category. Read the categories only if necessary. This question applies only to new participants*.]
#6 ego.occupation Currently, are you…?
_1. Employed / Self-employed
_2. Retired
_3. Unemployed
_4. Homemaker
_5. Student
[-99.] Don't know / No response [Note: do not read]
[Closed response. Nominal format with a single response option.
Note: *If simultaneously employed and a student/retired, prioritize the employed category*.]

#7 ego.single Currently, are you in a relationship?
_1. Yes [I have a spouse // a stable partner]
_2. No
[-99.] Don't know / No response [Operator: Do not read]
[Closed response. Nominal format with a single response option.
Note: *Read the options and categorize the response accordingly. If necessary, adapt the form of response options based on the respondent's gender*.]

#8 ego.household.sz In total, including yourself, how many people live with you (spouse, children, parents, other relatives, friends, etc.)?

_____
[Open response. Numeric format.
Note: *Assist the respondent in making the total, if needed. Write the total number of people including the respondent. Do not record anything if they do not wish to respond*.]

#9 ego.income Currently, your personal income in a regular month is … :
_1. Under 500 RON (≈ 100 EUR; 1 EUR ≈ 5 RON)
_2. Between 500 and 1000 RON (100-200 EUR)
_3. Between 1001 and 1500 RON (201-300 EUR)
_4. Between 1501 and 2000 RON (301-400 EUR)
_5. Between 2001 and 2500 RON (401-500 EUR)
_6. Between 2501 and 3000 RON (501-600 EUR)
_7. Between 3001 and 3500 RON (601-700 EUR)
_8. Between 3501 and 4000 RON (701-800 EUR)
_9. Between 4001 and 5000 RON (801-1000 EUR)
_10. Between 5001 and 6000 RON (1001-1200 EUR)
_11. Between 6001 and 7000 RON (1201-1400 EUR)
_12. Between 7001 and 8000 RON (1401-1600 EUR)
_13. Between 8001 and 9000 RON (1601-1800 EUR)

_14. Between 9001 and 10000 RON (1801-2000 EUR)
_15. Over 10000 RON (Over 2000 EUR)
[-99.] Don't know / No response [Note: Do not read]

[Closed response. Ordinal format.
Note: *Do not read the options. Let the respondent provide an open answer and place it in the appropriate category. Read answer options 1 to 15 only if the respondent is hesitant to provide a specific answer. Convert to common (popular) monetary expressions if needed.*]

#10 ego.income.house Currently, the total household income in a regular month is…?:
_1. Under 500 RON (≈ 100 EUR; 1 EUR ≈ 5 RON)
_2. Between 500 and 1000 RON (100-200 EUR)
_3. Between 1001 and 1500 RON (201-300 EUR)
_4. Between 1501 and 2000 RON (301-400 EUR)
_5. Between 2001 and 2500 RON (401-500 EUR)
_6. Between 2501 and 3000 RON (501-600 EUR)
_7. Between 3001 and 3500 RON (601-700 EUR)
_8. Between 3501 and 4000 RON (701-800 EUR)
_9. Between 4001 and 5000 RON (801-1000 EUR)
_10. Between 5001 and 6000 RON (1001-1200 EUR)
_11. Between 6001 and 7000 RON (1201-1400 EUR)
_12. Between 7001 and 8000 RON (1401-1600 EUR)
_13. Between 8001 and 9000 RON (1601-1800 EUR)
_14. Between 9001 and 10000 RON (1801-2000 EUR)
_15. Over 10000 RON (Over 2000 EUR)
[-99.] Don't know / No response [Note: Do not read]
[Closed response. Ordinal format.
Note: *Do not read the options. Let the respondent provide an open answer and place it in the appropriate category. Read answer options 1 to 15 only if the respondent is hesitant to provide a specific answer. Convert to common (popular) monetary expressions if needed.]*

I.2 Perceptions regarding health status

#11 ego.self.health On a scale from 1 to 10, where 1 means *very poor* and 10 *very good*, how would you rate your health status…?

-99	1 (very poor)	2	3	4	5	6	7	8	9	10 (very good)

[Closed response. Ordinal format aided by a visual scale.]

#12 ego.medical.condition At this moment, do you know if you suffer from any illness? (Explicitly ask for the name(s) of the illness / illnesses, if applicable)

_____
[Open response. Text format.
Note: *Write "NO" if they say they do not have any illness or do not know. Do not record anything if they do not wish to respond. Explore carefully if the respondent is an oncology patient or has cardiovascular*

conditions. Record the opinion of the consulting physician, if the respondent remembers. Emphasize "at this moment".]

#13 ego.weight At this moment, what is your weight?

_____

[Open response. Text format.

Note: *Write the response in kilograms and grams – e.g. 60; 70.5 (if they say 70 kilograms and half). Do not record anything if they do not wish to respond. For pregnant women, record the weight before pregnancy.]*

#14 ego.height What is your height?

_____

[Open response. Text format.

Note: *Write the height using two decimals – e.g., 1.75. Do not record anything if they do not wish to respond.*]

#15 ego.food.breakfast Please tell me what you usually eat for breakfast:

_____

[Open response. Text format.

Note: *Request detailed information about the combination of foods / dishes. Also ask for details about: liquids intake (e.g., water, coffee, tea, milk, alcohol), bread consumption, fruit and vegetable consumption, salt intake, sugar or sweets consumption.*]

#16 ego.food.lunch Please tell me what you usually eat for lunch:

_____

[Open response. Text format.

Note: *Request detailed information about the combination of foods / dishes. Also ask for details about: liquids intake (e.g., water, coffee, tea, milk, alcohol), bread consumption, fruit and vegetable consumption, salt intake, sugar or sweets consumption.*]

#17 ego.food.dinner Please tell me what you usually eat for dinner:

_____

[Open response. Text format.

Note: *Request detailed information about the combination of foods / dishes. Also ask for details about: liquids intake (e.g., water, coffee, tea, milk, alcohol), bread consumption, fruit and vegetable consumption, salt intake, sugar or sweets consumption.*]

#18 ego.meals.time On a regular day (Monday to Friday), how many times do you eat and when?

_____

[Open response. Text format.

Note: *For each mentioned meal, note the approximate time. Also, ask for details about eating between meals.*]

#19 ego.cooking.person Who cooks most often in your household?

_____

[Open response. Text format.

Note: *Ask for the name of the person.*]

#20 ego.cooking.list Who makes the shopping list for cooking and other groceries bought in the household?

_____

[Open response. Text format.
Note: *Ask for the name of the person.*]

#21 ego.food.place Typically, where do you buy the groceries you need for cooking?

_____

[Open response. Text format.
Note: *Emphasize location and type of store: village, neighbors, town, supermarket, hypermarket, market, etc.*]

#22 ego.food.income On a scale from 1 to 10, where 1 means *not at all* and 10 means *completely*, to what extent does the income available in your household give you the possibility to eat what you want?

-99	*1 (not at all)*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10 (completely)*

[Closed response. Ordinal format with visual scale.]

#23 ego.pfood Do you consume something from this list on a daily or weekly basis…? (purchased from the store:) Salami, ham, baloney, sausages, hot dogs, liver pate, pizza, popcorn, chips, canned goods, burgers, kebabs?
_1. Yes
_0. No
[-99.] Don't know / No response [Note: Do not read]
[Closed response. Binary format.
Note: *Mention each item from the list, as they are ordered, and stop when they mention an item they consume daily or weekly.*]

[Filter: if <ego.pfood> = 1]
#24 ego.pfood.motive What is the main reason you eat foods like the ones mentioned earlier?
_1. They are affordable
_2. I like their taste
_3. It's easy for me to buy them
_4. I don't have anything else to buy from the stores I go to
_5. To cook less
_6. Because I don't cook
_7. Other
[-98.] Not applicable (does not consume) [if <ego.pfood> = 0]
[-99.] Don't know / No response [Operator: Do not read]
[Closed response. Nominal format with a single response option.]

#25 ego.diet Are you currently on a diet or following any dietary regimen?
_0. No
_1. Yes, based on a doctor's recommendation
_2. Yes, based on a personal choice

[-99.] Don't know / No response [Note: Do not read]
[Closed response. Nominal format with a single response option.]

#26 ego.sport Are you regularly engaging in any sports (daily or weekly)?
_1. Yes
_0. No
[-99.] Don't know / No response [Note: Do not read]
[Closed response. Binary format.
Note: *Only consider physical activities explicitly practiced as a sport. Exclude activities such as walking, hiking, fishing, chess, etc.*]

I.3 Lifestyle

#27 ego.smoker Currently, you are …
_1. A smoker
_2. An occasional smoker (e.g., at events)
_3. A former smoker
_4. I have smoked very few cigarettes to consider myself a smoker
_5. I have never smoked
[-99.] Don't know / No response [Operator: Do not read]
[Closed response. Nominal format with a single response option.]
[Filter: if <ego.smoker> = 1 or 2]

#28 ego.smoke.time When do you smoke your first cigarette?
_1. Before the first meal of the day
_2. After the first meal of the day
[-98.] Not applicable (non-smoker) [if < ego.smoker > = 3, 4, or 5]
[-99.] Don't know / No response [Note: Do not read]
[Closed response. Binary format.]

#29 ego.sleep On a regular weekday (Monday to Friday), what time do you usually go to bed?
_____
[Open response. Text format.
Note: *Write in HH:MM format. Use the 24-hour time format. Do not record anything if they choose not to respond.*]

#30 ego.wake On a regular weekday (Monday to Friday), what time do you usually wake up?
_____
[Open response. Text format.
Note: *Write in HH:MM format. Use the 24-hour time format. Do not record anything if they choose not to respond.*]

#31 **ego.hf.perception** On a scale from 1 to 10, where 1 means *not healthy at all* and 10 *very healthy*, how healthy do you think you eat?

-99	*1 (not at all)*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10 (very healthy)*

[Closed response. Ordinal format with visual scale.]

II. Data about alters (ego's social contacts)

II.1 Alter generators

Instructions for the respondent: For the following part of the questionnaire, please nominate individuals who are part of your life, and whom you see frequently. Although I ask you to nominate them, their identity will remain confidential and will then be deleted from the database. We will start with family members and relatives. Then we will continue with close friends and acquaintances.

[**Instructions for the interviewer**: Note the surname(s) first, followed by the first name(s). **For household members, all mentioned individuals are considered, regardless of age**. For the rest of the generators, only ask for alters who are over 18 years old.]

#32 **alter.same.hh** Please mention the persons who are part of your household:

[Filter: if <ego.occupation> = 1]

#33 **alter.coworker** Please mention coworkers with whom you meet on a daily or weekly basis:

[**Note:** *Consider only face-to-face encounters*.]

#34 Please mention individuals with whom you interact daily or at least weekly. They can be persons: with whom you meet for a barbecue, with whom you attend community events, who are your neighbors, who you visit, or who visit you, etc.:

[**Note:** *Consider only family members who have not been mentioned previously*.]

[**Note:** *Consider only face-to-face encounters*.]

> #34A **alter.family.freq** I suggest we start with family members (or relatives) other than those you live with:
>
> #34B **alter.friend.freq** Let's continue with friends (persons who are not your relatives and others besides those mentioned as coworkers):

II.2 Alter interpreter

For each of the persons you have mentioned, please specify:

#35 **alter.food.companion** If you frequently eat together with this person (daily or weekly):

[*Alter enumeration****]***

_1. Yes

_0. No

[Closed-ended question. Binary format.]

[Filter: if <alter.same.hh > = TRUE]

#36 **alter.food.special** if special (different) meals are cooked for this household member:

[Alter enumeration]

_1. Yes

_0. No

[Closed-ended question. Binary format.]

#37 **alter.sport** if they regularly practice a specific sport (daily or weekly):

[Alter enumeration]

_1. Yes

_0. No

[-99.] I don't know / No answer [Note: Do not read]

[Closed-ended question. Binary format.

Note: *Only consider explicitly physical activities practiced as a sport. Exclude activities like walking, strolling, fishing, chess, etc.*]

#38 **alter.diet** if they are on a special diet or dietary regimen:

[Alter enumeration]

_0. No

_1. Yes, based on a doctor's recommendation

_2. Yes, based on a personal choice

[-99.] Don't know / No response [Note: Do not read]

[Closed-ended question. Nominal format with only one response option.]

#39 **alter.smoking** if currently they are a:

[Alter enumeration]

_1. Non-smoker (never smoked)

_2. Former smoker

_3. Smoker

[-99.] I don't know / No answer [Note: Do not read]

[Closed-ended question. Nominal format with only one response option.

Note: *Smokers include individuals who smoke regularly or occasionally.*]

#40 **alter.pfood** if they frequently consume (daily or weekly) salami, ham, baloney, sausages , hot dogs, pâté, pizza, popcorn, chips, canned foods, burgers, kebabs (any of them):

[Alter enumeration]

_1. Yes

_0. No

[-99.] I don't know / No answer [Note: Do not read]

[Closed-ended question. Binary format.]

#41 alter.cvd if they were diagnosed by a doctor with a heart disease [in the past / already cured, or currently]:
[Alter enumeration]
_1. Yes
_0. No
[Closed-ended question. Binary format.]

#42 alter.onco if they were diagnosed by a doctor with a type of cancer [in the past / already cured, or currently]:
[Alter enumeration]
_1. Yes
_0. No
[Closed-ended question. Binary format.]

#43 alter.age Age (in years):
[Alter enumeration]

_____
[Open-ended question. Text format.]

#44 alter.edu Education level:
[Alter enumeration]
_1. Up to eight years of education [finished lower secondary education]
_2. Between nine years and post-secondary school/trade school [non-tertiary education]
_3. University studies (minimum Bachelor's degree)
[-99.] Don't know / No response [Note: do not read]
[Closed-ended question. Ordinal format.]

#45 alter.sex Sex:
[Alter enumeration]
_1. Male
_2. Female
[Closed-ended question. Binary format.
Note: *Do not read. Automatically classify based on the first name(s). Request clarifications if not clear (name is ambiguous).*]

#46 alter.occupation If, currently, they are:
[Alter enumeration]
_1. Employed / Self-employed
_2. Retired
_3. Unemployed
_4. Homemaker
_5. Student
[-99.] I don't know / No answer [Note: do not read]
[Closed-ended question. Nominal format with a single response option.
Note: If simultaneously employed and student / retired, prioritize the employed category.]

#47 **alter.bmi** Please take a look at the following two showcards depicting different body types for men and women. Among the categories you see, please associate the individuals mentioned earlier with a body type.
[**Alter enumeration**]
_*1. A*
_*2. B*
_*3. C*
_*4. D*
_*5. E*
_*6. F*
_*7. G*
_*8. H*
_*9. I*
_*10. J*
[-99.] I don't know / No answer [Note: Do not read]
[Closed-ended question per alter. Ordinal format.
Note: *Ask about each alter and place them in the corresponding category. Very young children who cannot be placed on the visual scale fall into the category -99. DK / NA*]

II.3. Alter-Alter ties

#48 **aa.ties** Please tell me how often these persons interact (face-to-face meetings):
[*List of dyads*]
_*1. They don't know each other*
_*2. Less often (than once a week)*
_*3. At least once a week*
[Closed-ended response per dyad. Nominal format with only one response option.]

III. Perceptions about healthy eating

#49 **food.healthy.breakfast** Imagine someone is visiting you and you need to prepare a healthy breakfast for them. What do you think they should eat?

_____
[Open-ended response. Text format.
Note: *Ask for a detailed response about the combination of foods / dishes / beverages.*]

#50 **food.healthy.lunch** Imagine someone is visiting you and you need to prepare a healthy lunch for them. What do you think they should eat?

_____
[Open-ended response. Text format.
Note: *Ask for a detailed response about the combination of foods / dishes / beverages.*]

#51 food.healthy.dinner Imagine someone is visiting you and you need to prepare a healthy dinner for them. What do you think they should eat?

_____

[Open-ended response. Text format.
Note: *Ask for a detailed response about the combination of foods / dishes / beverages.*]

IV. 4P-CAN Events participation

#52 ego.event Please specify three ideas that you found interesting after participating at the 4P-CAN events:

_____
_____
_____

#53 event.coparticipation Please tell me who you came with to the events organized within this project:

_____

[Open-ended response. Text format.
Note: *Ask for the names of the people they came with.*]

V. End of interview data

#54 referral Could you help us contact at least five people (your family members, friends, or acquaintances) who are at least 18 years old and currently living in [NAME OF PLACE], to participate in our study? For this purpose, we would need their name and phone number. We kindly ask you to contact them initially.

	Family name(s) and surname(s) (name)	**Phone number (con_phone)**
REF 1		
REF 2		
REF 3		
REF 4		
REF 5		

fin_obs Other observations:

_____

[Open-ended response. Text format.
Note: *Write down observations that can be important (e.g., what the respondent says about calling them to see if they talked with their contacts). Do not record anything if it's not the case.]*

[Showcards used in the second wave]

Figure 9 Showcard with different types of bodies for females
Note: The showcard is extracted from the work of Harris et al.[39]

Weight by Measured BMI	Under weight	Normal weight	Over weight	Class I Obesity	Class II Obesity	Class III Obesity

Figure 10 Showcard with different types of bodies for females grouped by BMI categories
Note: The showcard is extracted from the work of Harris et al.[39]

Figure 11 Showcard with different types of bodies for males
Note: The showcard is extracted from the work of Harris et al.[39]

Weight by Measured BMI	Under weight	Normal weight	Over weight	Class I Obesity	Class II Obesity	Class III Obesity

Figure 12 Showcard with different types of bodies for males grouped by BMI categories
Note: The showcard is extracted from the work of Harris et al.[39]