Impact of COVID-19 on the health of the general and more vulnerable population and its determinants: Health care and social survey-ESSOC, study protocol

Carmen Sánchez-Cantalejo
Andalusian School of Public Health: Escuela Andaluza de Salud Publica
Mª del Mar Rueda
University of Granada: Universidad de Granada
Marc Saez
University of Girona: Universitat de Girona
Iria Enrique
Andalusian Institute of Statistics and Cartography
Ramón Ferri-García
University of Granada: Universidad de Granada
Miguel De la Fuente
Demométrica
Román Villegas
Government of Andalusia Andalusian Health Service: Junta de Andalucia Servicio Andaluz de Salud
Luis Castro
University of Granada: Universidad de Granada
Maria Antònia Barceló
University of Girona: Universitat de Girona
Antonio Daponte-Codina
Andalusian School of Public Health: Escuela Andaluza de Salud Publica
Nicola Lorusso
Health surveillance: Vigilancia de la Salud, Government of Andalusia
Andrés Cabrera-León (andres.cabrera.easp@juntadeandalucia.es)
Andalusian School of Public Health: Escuela Andaluza de Salud Publica https://orcid.org/0000-0002-4812-1026

Method Article

Keywords: public health, health determinants, health inequalities, COVID-19, SARS-CoV-2, vulnerable populations, real-world data, surveys, population registries, machine learning

Posted Date: June 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-505722/v3

Version of Record: A version of this preprint was published at International Journal of Environmental Research and Public Health on July 31st, 2021. See the published version at https://doi.org/10.3390/ijerph18158120.

Abstract

This manuscript describes the rationale and protocol of a real-world data (RWD) study entitled *Health Care and Social Survey (ESSOC*, *Encuesta Sanitaria y Social*). The study's objective is to determine the magnitude, characteristics, and evolution of the COVID-19 impact on overall health as well as the socioeconomic, psychosocial, behavioural, occupational, environmental, and clinical determinants of both the general and more vulnerable population. The study integrates observational data collected through a survey using a probabilistic, overlapping panel design, and data from clinical, epidemiological, demographic, and environmental registries. The data will be analysed using advanced statistical, sampling, and machine learning techniques. The study is based on several measurements obtained from three random samples of the Andalusian (Spain) population: general population aged 16 years and over, residents of disadvantaged areas, and people over the age of 55. Given the current characteristics of this pandemic and its future repercussions, this project will generate relevant information on a regular basis, commencing from the beginning of the State of Alarm. It will also establish institutional alliances of great social value, explore and apply powerful and novel methodologies, and produce large, integrated, high-quality and open-access databases. The information described here will be vital for health systems in order to design tailor-made interventions aimed at improving the health care, health, and quality of life of the populations most affected by the COVID-19 pandemic.

1. Introduction

1.1. Background

Currently underway are a large number of studies investigating the evolution of the 2019 coronavirus disease (COVID-19) and the impact it is having on the number of infected patients, hospital admissions, and deaths[1–6], as well as on the mental health and well-being of the population[7]. Nevertheless, very few reports providing information concerning clinical, contextual and citizens' general perceptions pertaining to the pandemic's impact and evolution since its onset are being prepared[8–10].

The Spanish Government officially declared a State of Alarm on 14 March, 2020 (Spanish Royal Decree 463/2020[11]) in the face of the global public health emergency caused by COVID-19. Among other actions, it ordered individuals' freedom of movement to be limited (Article 7, "Limitación de la libertad de circulación de las personas"), which was subsequently further restricted through other decrees (hereinafter referred to as confinement). These limitations have led to a whole series of - still little-studied - problems in the population. In this sense, recent reviews concerning epidemic outbreaks and subsequent confinements and have concluded that these actions have very negative and long-term impacts on mental health[12,13]. Likewise, this current pandemic has also seen such negative effects, albeit in the short-term[14–17]. Furthermore, all studies agree on the urgent need for more evidence regarding this kind of impact, especially evidence gathered from among the most exposed populations as well as those populations in a situation of greater vulnerability[18]. In this respect, a study on the Roma population observed that 14% found it difficult to follow the measures aimed at preventing the disease and 7% experienced difficulty accessing medicines[19]. As a result, a deterioration in the social and educational gap was observed in this population, along with them experiencing greater stigmatization owing to certain lifestyle habits[20]. Likewise, a study carried out with minors (infants and adolescents) in foster care revealed a decrease in their quality of life during the pandemic compared with the year 2017[21],while another study involving vulnerable populations demonstrated the existence of a link between experiencing anxiety and the perceived risk of contracting the virus[8].

As for the impact of COVID-19 on the migrant population and ethnic minorities, the results of different studies have revealed a higher incidence of the disease in these population groups in relation to deficient conditions in terms of socioeconomic deprivation, comorbidities, unemployment, and livability[22-24]. As in the case of the Roma population, this situation may be aggravated in the migrant population as they have poorer access to health care because of the deterioration of their socioeconomic and administrative conditions or their difficulty understanding the prevention and health care information provided due to language limitations[25,26].

Regarding the effects of pandemics on the health of people with chronic diseases, and also in relation to the above, the results of some studies have highlighted the need to adapt health care to the confinement situation by introducing measures such as remote consultations[27,28]. Furthermore, these studies revealed that throughout the confinement period the symptoms of some diseases, such as Alzheimer's[29] or diabetes[30], were exacerbated. That said, this worsening of symptoms was not as severe for patients with chronic obstructive pulmonary disease (COPD)[31]. It is well known that health care can be affected during health emergencies and pandemics as a result of resources being diverted towards more urgent areas and, for example, the care of patients with chronic diseases being restricted and/or surgical procedures being delayed[32]. However, the evidence available indicates that ensuring proper follow-up of patients with chronic diseases should be one of the main goals when managing the COVID-19 pandemic because, if infected, they are at risk of experiencing greater severity[33,34].

With respect to the health impact of COVID-19 in Spain, several studies have been launched, most of which are using transversal designs, non-probabilistic samples, and web surveys[35,36]. These studies represent an agile and simple alternative for collecting a large amount of data quickly and offering results practically in real time; something which is crucial to responding to situations like the one currently being experienced. However, these surveys have a number of disadvantages and limitations. For instance, they do not allow for the impact on and evolution in certain populations in more vulnerable positions to be determined, nor do they allow valid estimates to be obtained which would, in turn, allow the level of error to be limited to the appropriate standards to be conclusive. The reason for this is because they are not usually processed with a view to solving problems resulting from a lack of coverage and response, or from selection bias in the sample[37].

At a national level, the National Epidemiology Center (CNE, Centro Nacional de Epidemiología) launched the National Seroepidemiology Study (Estudio Nacional de Seroepidemiología) based on the data obtained through a population-based macro-survey that provides estimates down to the provincial

level, continuous information on the evolution of the epidemic, and information obtained from patients' medical records[38,39]. Based on its preliminary results, the study has estimated the prevalence of immunoglobulin G (IgG), antibodies against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) to be at 9.9% (95% confidence interval [CI]: 9.4%–10.4%) in Spain, with a very similar prevalence among men and women. In addition, while the coastal areas have prevalences close to or lower than 4%, the central core surrounding Madrid has figures close to or above 15%, with the highest prevalence (>16%) being detected among health care personnel and women caring for dependent persons[40].

COVID-19 has had the greatest impact on people over 70 years of age and who account for 38% of confirmed COVID-19 cases in Spain, almost half of the hospitalizations due to the virus, and more than 85% of the deaths[41]. Another study has shown that fatality from COVID-19 among individuals over the age of 80 ranges between 12% and 16% in men and 4.5% and 6.5% in women[39]. Internationally, it has been confirmed that the elderly are at greater risk of hospitalization[42] and developing severe symptoms[43], with people over 60 years of age and those with underlying conditions such as hypertension, diabetes, cardiovascular diseases, chronic respiratory diseases, or immunosuppression, being at higher risk of severe infection and death[44]. Smokers and ex-smokers over the age of 60 are also at greater risk and more susceptible to becoming infected by SARS-CoV-2[45]. These extreme pandemic-related incidence and mortality rates among the elderly population are further aggravated by its psychosocial impact[46], with a higher risk of depressive feelings among these individuals who have mostly had little physical contact with other people during the confinement period[47,48]. This is a clear reflection of the tragic COVID-19 situation that is currently being experienced, and may be even harsher in deprived socioeconomic contexts[9], although these results must be further explored and there is no evidence in this respect in our context. Furthermore, the situation caused by the pandemic means that coordination between social and health care services, together with the provision of resources, are even more necessary and urgent than ever[49].

1.2. The ESSOC Study Framework

The Health Care and Social Survey (ESSOC, Encuesta Sanitaria y SOCial) research project arises from the need to provide specific, reliable and timely data on the impact of COVID-19, which can then be considered when making decisions to prepare and provide an effective Public Health response in the different populations affected, especially the most vulnerable such as, among others, the elderly, the chronically ill, or those at risk of exclusion. ESSOC focuses on Public Health[50] and relies on the participation of society to obtain information on people's health and their quality of life in order to be able to intervene both individually and collectively in the face of the pandemic (Table 1). The research project is based on a real-world data[51] design that integrates observational data extracted from multiple sources of differing natures and perspectives, i.e., both probabilistic and administrative. Thus, it will generate a very large amount of linked data from differing sources based on longitudinal probabilistic population surveys, (more than 22,000 interviews with information on more than 700 variables over 3 years on the economic situation, state of health and well-being from the perspective of citizens), and sources based on clinical (more than one million records between COVID and non-COVID cases with information on more than 110 variables over two years), epidemiological (almost 600,000 cases with information on more than 25 variables over 1 year), demographic (from the 8.5 million people residing in Andalusia) and environmental records (more than a million records with information on more than 20 variables over 10 years), (i.e., data from an administrative perspective).

In addition, the data analyses will be performed using advanced statistical, sampling and machine learning techniques, which allow for new research methods to be developed and implemented[52]. The research follows an Open Science approach[53] in terms of disseminating its results, methodologies, processes, and collected data, which will be distributed, reused, and freely and openly accessible to not only the scientific, academic, clinical, and public health managerial community, but also to society at large and, in particular, those population groups identified as being at greater risk of vulnerability to COVID-19. Its management model is that of collaborative and multidisciplinary research, facilitating the creation of a context of scientific cooperation in the fields of public health, health care, public administration, data science, environmental and demographic sciences, and social sciences (Table 1). Finally, the study region, Andalusia, is the most populated (8.5 million inhabitants) and the second largest in area of the 19 regions in Spain. It is also the fifth most populated region in Europe, and it is as populated as other European countries such as Austria or Switzerland. The ESSOC research project is included in the Oxford Supertracker global directory of policy trackers and surveys related to COVID-19[1].

Table 1. Health Care and Social Survey (ESSOC): Study Framework

FOCUS	DESIGN	DISSEMINATION	MANAGEMENT
Public Health	Real-World Data	Open Science	Collaborative and Multidisciplinary Research
Community participation to obtain information on people's health and quality of life to be able to intervene both individually and collectively in the face of the pandemic	Integration of observational data extracted from multiple sources from different perspectives based on probabilistic samples and administrative registries	Results, methodologies, processes, and collected data distributed, reused, and freely and openly accessible	Andalusian School of Public Health, Andalusian Health Service, Department of Health and Families (Andalusian Regional Government), Universities of Granada and Girona, Andalusian Institute of Statistics and Cartography and Guadalinfo Network

1.3. Hypotheses

- Perceptions of general health, mental health, and emotional well-being have deteriorated in the short- and mid-term since the beginning of the pandemic, with a greater impact being observed in women, young people, and those diagnosed with COVID-19.
- The socioeconomic, psychosocial, behavioural, occupational, environmental, and clinical determinants of health have deteriorated since the onset of the pandemic.

- Health inequalities have increased along the axes of social class, gender, age, ethnicity, and territory as a result of COVID-19, and have been even greater in the mid-term compared with the short-term.
- Chronicity and resulting disability have increased significantly since the beginning of the pandemic.
- Since the beginning of the pandemic, the care burden has increased significantly for women in the short-term and this has had a highly negative impact on their health and well-being.
- Social and emotional support in the population aged over 55 years has decreased significantly since the beginning of the pandemic, with the greatest differences being observed in single-person households in urban areas.

1.4. Objectives

General:

To determine the magnitude, characteristics, and evolution of the impact of COVID-19 on overall health and its socioeconomic, psychosocial, behavioural, occupational, environmental, and clinical determinants in the general population and that with greater socioeconomic vulnerability.

Specific:

1. To determine the short- and mid-term impact of the COVID-19 pandemic on the health and emotional well-being of the general population of Andalusia.
2. To analyse the evolution of the socioeconomic, psychosocial, behavioural, clinical, and environmental determinants of health in the context of the COVID-19 pandemic in the population under study.
3. To identify health inequalities along the axes of social class, gender, age, ethnicity, and territory and their evolution in the context of the COVID-19 pandemic.
4. To evaluate different research sampling techniques to improve the reliability and precision of estimates obtained through surveys using longitudinal designs.

To compile and systematize the existing evidence regarding the design, sources, methodologies, and topics related to measuring the impact of COVID-19 on health and its determinants through surveys.

2. Materials And Methods

2.1. Study Design

This study employs real-world data design to integrate observational data extracted from multiple sources, including information obtained from different providers based on surveys and clinical, epidemiological, population, and environmental registries. The surveys have an overlapping panel design to ensure there are both cross-sectional and longitudinal estimates [54] and to include population-based probability samples carried out via telephone interviews.

2.2. Geographical, Population, and Temporal Scopes

The geographical scope is the Autonomous Region of Andalusia (*Comunidad Autónoma de Andalucía*), Spain, and the population scopes are the general population over the age of 16 (ESSOCgeneral), the population residing in disadvantaged areas (ESSOCzones)[55], and the population over the age of 55 (ESSOC+55). Collective households (i.e., hospitals, nursing homes, barracks, etc.) are not considered in this study. That said, the study sample includes families who, as an independent group, reside in these collective establishments (e.g., a director or janitor of a centre). The temporal scopes of each sample (Figure 1) are:

- ESSOCgeneral: five measurements taken between 2020 and 2023, at baseline (beginning of the Spanish State of Alarm), at one month from the first interview, at six months, at 12 months, and at 36 months.

- ESSOCzones: two measurements, taken at baseline (12 months from the beginning of the State of Alarm) and at 24 months from the first interview.

- ESSOC+55: two measurements, taken at baseline (six months from the beginning of the State of Alarm) and at 30 months from the first interview.

2.3. Sampling Frame

The sampling frame used to extract the ESSOCgeneral and ESSOCzones samples is obtained from the Longitudinal Andalusian Population Database (BDLPA, *Base Longitudinal de Datos de Población de Andalucía*)[56] . The information consolidated in the BDLPA originates from the integration of information on stocks, flows, and variations extracted from the census coordination system which, together with the data obtained from the Civil Registries with respect to births, deaths, and marriages (i.e., vital statistics [MNP, *Movimiento Natural de la Población*]), as well as that reported in the population and housing censuses, give rise to an integrated longitudinal frame for population and territorial statistics in Andalusia[57]. The sampling frame is extracted from the BDLPA longitudinal file as a cross-section with the reference date set as 01 January of 2019. The selected samples are linked to the information obtained from the User Database (BDU, *Base de Datos de Usuarios*)[58] of the Andalusian Public Health Care System in order to obtain the telephone numbers of the selected sample units. The BDU coverage in terms of contact telephone numbers for the selected samples is usually above

96%. On the other hand, the ESSOC+55 sampling frame corresponds to the user population of the Andalusian Guadalinfo public network aged 55 years and over [59].

2.4. Sample Size

During the first measurement (M1), the ESSOCgeneral scope started with a random sample comprised of 5,000 people under the assumptions of maximum variability in the estimate (p = q = 0.5), a design effect of 1.8, a precision of 2.4 percentage points for estimates in Andalusia, a confidence level of 95%, and a non-response rate for the theoretical sample of 40%. The subsequent measurements (M2–M5) are comprised of the longitudinal samples of the previous measurements (n_i_theoretical_longitudinal) and, in addition, of a new sample in each measurement (n_i_theoretical_new). That new sample is selected according to the design of the first measurement, except for M5 which will incorporate a new stratum of 'residing or not in disadvantaged areas', where DA1 = non-disadvantaged area and DA2 = disadvantaged area[60]. For the previous measurements (M1, M2, M3 and M4), a post-stratification will be carried out according to disadvantaged areas in order to improve the estimates. Finally, due to non-response, from these two theoretical samples (longitudinal+new) we obtain the total effective sample for these measurements (n_i_effective = n_i_effective_longitudinal + n_i_effective_new). With respect to the theorical sample size for the longitudinal sample of i measurement, this is defined by n_i_theoretical_longitudinal = n_i-1_effective_longitudinal + n_i-1_effective_new, i=2…5). The aim is to reach an effective sample of 3,000 units per measurement for ESSOCgeneral, 2,750 for ESSOCzones and at least 2,400 for ESSOC+55, assuming a total response rate of 60%.

Thus, the ESSOC is made up of a series of measurements broken down into a new sample and a longitudinal sample for each measurement which, in turn, are categorized into theoretical and effective samples for each population study group (general, zones and +55), with a total of over 22,000 effective interviews being carried out over three years.

2.5. Sample Allocation

Allocation of the new samples (including the first sample) is mixed. On the one hand, they are uniform by province (150 sample units for eight provinces) and, on the other hand, proportional to the population size of each province and degree of urbanization (urban, intermediate density, and rural area)[61]. In addition, for the M5 measurement, the distribution of the new sample will be performed, in the first place, in the two DA strata and, subsequently, as in the case of the previous measurements.

2.6. Sample Selection

The selection of the new theoretical sample in each measurement is carried out in a simple random manner within each province and degree of urbanization stratum and, in the case of measurement M5, within each DA strata, thus obtaining self-weighted samples in each stratum. The theoretical longitudinal sample of a measurement is composed of the effective sample of the previous measurement, except for measurement M1 which, being the first one, does not have a longitudinal sample.

In the ESSOC+55 sample, the first measurement was stratified by clusters (Guadalinfo Centers, N = 651), with sub-sampling to 1,200 users. These centres are stratified per Andalusian province and inhabitation level (<10,000 inhabitants, 10,000–19,999, ≥20,000), as well as sex and age quotas (55–64 years, 65–74 years, and >75 years). As in the case of the ESSOCgeneral sample, the ESSOC+55 second measurement will be made up of a longitudinal sample (that of the first measurement) and another new sample until a total of 1,200 interviews are completed.

2.7. Fieldwork

The survey information is collected through a computer-assisted telephone interview (CATI). The management, control, and monitoring of the collection of information for measurement M1 is carried out through Pl@teA, the IECA's survey collection platform, whereas for the rest of measurements these tasks are performed using the Mobinet-Gandia Integra software. The data collection is carried out by a team of between eight (in M1) and 12 (rest of ESSOCgeneral measures) interviewers assigned solely to this study. This ensures working team stability, which is of fundamental importance in regular, longitudinal surveys. Before starting the study, the interviewers will receive the necessary training regarding the content of the survey. To this end, in addition to virtual meetings held before starting the fieldwork, they are provided with interviewer's and questionnaire manuals which, besides explaining the questionnaire and study's content, describe the survey platform, possible incidences, and the protocol to be followed in each case to guarantee the maximum quality of the samples and the information collected. Prior to engaging in the fieldwork, each interviewer performs several pilot tests to measure times and determine the complicated points of the questionnaire.

The schedule set to conduct the surveys is from Monday to Friday from 10:00 to 21:00, and on Saturdays from 10:00 to 15:00 for the first measurement and for the rest of the measurements from Monday to Saturday from 14:00 to 21:00, although deferred appointments can be scheduled without time limits. Furthermore, a telephone line with a 900 prefix and staffed by telephone agents is made available to the public. This line is provided to the survey holders via text message or through the CSyF website, where the characteristics of this study are also published. This call centre also receives calls from people who, having been contacted by CATI agents, need to confirm the official nature of the survey. In fact, many of these calls become completed interviews.

2.8. Quality Control

For the ESSOCgeneral, quality control measures mean the interviews are cross-checked both internally in the Call Centre itself and in the IECA and EASP. Each interviewer is monitored to ensure that they follow the established protocols and that they use each type of incidence correctly. The intervals elapsed

between each call and their duration are also monitored.

In addition to recording the calls made by each interviewer, a listening check is also performed to review both the positive aspects and those to be reinforced in the supervised surveys (i.e., for 10% of the calls performed for the first measurement and 25% of the calls performed for the rest of the measurements). During these checks, aspects, such as the interviewer's self-presentation, their presentation of the study, the offer of being able to call the 900-prefix telephone number, the confirmation of the place of residence, the correct delivery of the questionnaire's questions and all response options, are assessed.

Quality control, data cleansing, and data coding are carried out simultaneously with the fieldwork with the aid of the software to be used in the study. Each interviewer is provided with a space on the platform to record observations during the survey being conducted. This then allows the supervision team to cleanse those interviews in which the interviewer detected an inconsistency in the respondent's answers, or those in which the interviewer made a mistake when completing the questionnaire. Likewise, the values of the variables are revised, and invalid ones cleansed. Moreover, the coding of the variables corresponding to open-ended questions, such as the respondent's occupation and educational level, is carried out in tandem with the fieldwork. In the rest of open-ended questions, prior to their coding, their possible answers are cleansed, and the categories deemed to correspond to the majority subsequently coded.

During the telephone surveys, different situations may arise that could result in the inability to complete the survey. This is known as field incidences (Table 2), with the most important types being final incidences, i.e., those that, after several attempts, finally result in the inability to complete the survey.

Table 2. Health Care and Social Survey (ESSOC): Interview incidences and protocol to be followed:

Incidence	Incidence	Description	Protocol
Frame incidence (reasons that make it impossible to complete the survey due to problems related to the sampling frame; for example, a telephone number with which to contact the sample person could not be obtained or the housing frame was not sufficiently up to date)	The telephone number does not exist	Wrong number: the telephone number dialled does not exist, corresponds to a fax, or has restricted calls.	Direct removal
	Not contactable	Out-of-date frame: a selected person is living in a different municipality, a telephone frame without a telephone number, a person unreachable through the telephone number/home address provided due to circumstances such as death, divorce/separation, etc.	Direct removal
Relationship-situation incidences (reasons that make it impossible to complete the survey due to several types of situations affecting the surveyed people, for instance, they cannot be located, they refuse to participate in the survey, or any other aspect that prevents the survey being conducted)	No contact	The household cannot be contacted (e.g., nobody answers the telephone, or the answering machine goes off)	Removal after four attempts performed on two different days, at two different times
	Absent	The selected person cannot be contacted	Removal after four attempts performed on two different days and at two different times
	Inability to answer	The selected person cannot complete the survey due to an inability to respond to it because of disability, age, illness, lack of knowledge of the language, or any other circumstance. If possible, the survey should be completed by a close relative.	Direct removal
	Refusal	The selected person refuses to complete the survey or refuses to continue it after it has begun.	Direct removal

As this is a longitudinal study, one of the most significant reasons for a lack of response is the potential interviewee identifies the incoming call number and does not answer the phone. To solve this, as much as possible, the telephone number from which each call is made is changed periodically, so that even if a number were to identified and blocked, we could continue to attempt to contact that person by employing a new telephone number.

In addition to the quality of the sample, there are other factors of interest in assessing how fieldwork has developed during a surveying operation. One such factor is to determine how q survey has been carried out in terms of effectiveness and efficiency. The most direct way to measure this is to calculate the number of attempts or calls that had to be made in order to complete each survey. This type of information is also very useful to be able to design strategies aimed at optimizing attempts and, therefore, increasing sample levels in future operations.

2.9. Sampling Weights

The original sampling weight for the new samples is obtained from the inverse values of the effective sampling fractions in each stratum and used to calculate the Hajek estimator[62]. This is subsequently calibrated to obtain more reliable estimates based on the demographic characteristics of Andalusia. To this end, we use a truncated linear calibration method[63], and, as auxiliary information, the marginals of the Andalusian population per sex and age (16–19, 20–24, 25–29,...,75–79, and ≥80 years old), sex and province, sex and nationality (Spanish or dual nationality and foreign), and sex and degree of urbanization having obtained these data from the Continuous Municipal Register (Padrón Continuo de Municipios)[64]. Regarding the non-response bias in longitudinal samples, we can predict non-contact and non-cooperation based on auxiliary information and information already known

about the sample subjects. Thus, the original weights used in the estimates of longitudinal sample M_t are corrected during a first phase by modelling the non-response with respect to the longitudinal effective sample obtained in M_t-1 using machine learning techniques[65]. Said non-response is estimated using a XGBoost model[66], which represents the state-of-the-art in machine learning. Every piece of data and variable from the M_t-1 sample is used for training, thus the algorithm has all the information available in order to learn. Likewise, the hyperparameters of the model are optimized using cross-validation to ensure reliable estimations. Then, during a second phase, they are calibrated following the method described for the new samples. As auxiliary variables, we use those extracted from the Continuous Municipal Register (e.g., nationality, sex, age, province, degree of urbanization, etc.) and the registers from the ESSOC itself in M_t-1 (Table 3).

Table 3. Health Care and Social Survey (ESSOC): Adjustment of the design sampling weight in each measurement

Sample type (effective)	Type of adjustment 1st phase	2nd phase
New	Non-response adjusted by proxy based on the effective sample size in each stratum.	Representativeness by truncated linear calibration with 0.1 and 10 limits based on the auxiliary variables
Longitudinal	Non-response adjusted using an XGBoost model based on variables from the previous measurement.	

In addition to these adjustments, other methodological alternatives, not yet explored in this type of sample design, for instance double calibration, will also be investigated by considering different variables in order to model non-responses and, on the other hand, correction of the representativeness bias[67,68] and machine learning techniques, and adjusting non-responses with the aid of the Propensity Score Adjustment (PSA)[69,70].

2.10. Variables

The study variables will mainly be extracted from the following sources: BDLPA; the Andalusian Population Health Database (BPS, *Base Poblacional de Salud*)[71]; the Andalusian Environmental Information Network (REDIAM)[72]; the Andalusian Epidemiological Surveillance System (SVEA)[73] and the ESSOCgeneral, ESSOCzones, and ESSOC+55 surveys.

The personal data of the participants selected for the interview (name, surname, and telephone number) are extracted from the BDLPA. In addition, the BDLPA is linked annually with a repository of georeferenced buildings so that the postal address and coordinates (250m x 250m grid) in the territory can be extracted. This will allow us to extract geographical factors (urbanization degree and population density, among others) via other IECA registries, and environmental factors (pollution and temperature, among others) via the REDIAM registry from the Andalusian Regional Government's Department of Agriculture, Livestock, Fisheries, and Sustainable Development (*Consejería de Agricultura, Ganadería, Pesca y Desarrollo Sostenible de la Junta de Andalucía*).

From the SVEA registry, epidemiological information related to COVID-19, such as the date and result of the diagnostic test for active infection (PDIA), will be extracted to detect the presence of an active SARSCoV-2 infection, which includes both reverse transcription– polymerase chain reaction (real time RT – PCR) as the antigen (Ag) rapid test; date of the onset of symptoms; close contact of confirmed case with PDIA; local or imported case; occupation as a health or social health professional; need for hospitalization or admission to an intensive care unit; date of admission and discharge.

In addition, the clinical information related to chronic diseases[74], functional and cognitive assessments, health resources (volume and cost), population stratification, and drugs consumed, which is obtained from the BPS, will also be added to the valid samples (Table 4). Further information about the variables and the main features of the abovementioned registers can be found in Supplementary Material 3.

Table 4. Health Care and Social Survey (ESSOC): Auxiliary sources and variables

Registry	Description	Information	Variables extracted
BDLPA – Longitudinal Andalusian Population Database[71]	Information from the census coordination system and civil registries that give rise to a consolidated framework of the Andalusian population	Personal data	Name, surname, identification health number (NUHSA), geographical coordinates
BDU – User Database of the Andalusian Public Health Care System[58]	Contact Information of the Andalusian Public Health Care System	Personal contact information	Telephone numbers
BPS - Andalusian Population Health Database[71]	Personal health information from the Andalusian Population Health Database and Healthcare information	Health and healthcare information	Chronic diseases, functional and cognitive assessments, health resources (volume and cost), population stratification, drugs consumed
REDIAM - Andalusian Environmental Information Network[72]	Daily averages by collecting/meteorological station and at the census section level	Pollution, temperature	Mean daily values from pollution, air quality and temperature
SVEA - Andalusian Epidemiological Surveillance System[73]	Functional organization for health surveillance that collects, among other things, epidemiological information related to SARS-COV-2 infection	Epidemiological information of COVID-19	PCR result, symptoms date, close contact, healthcare professional, hospitalization unit (specifying ICU), date of admission and discharge, need of mechanical ventilation and clinical data

For further details, see Supplementary Material 3.

With regard to the surveys, each measurement is associated with a questionnaire that coincides, to a significant extent, with previous measurements to enable an analysis of its evolution, and incorporate new information to analyse specific characteristics present at each moment of the pandemic.

Repeating unchanging information in subsequent measurements is avoided in the case of longitudinal samples. The questionnaire used for each measurement is organized into blocks of information, as shown in Table 5.

Table 5. ESSOC: Auxiliary sources and variables

Registry	Description	Information	Variables extracted
BDLPA – Longitudinal Andalusian Population Database[71]	Information from the census coordination system and civil registries that give rise to a consolidated framework of the Andalusian population	Personal data	Name, surname, identification health number (NUHSA), geographical coordinates
BDU – User Database of the Andalusian Public Health Care System[58]	Contact Information of the Andalusian Public Health Care System	Personal contact information	Telephone numbers
BPS - Andalusian Population Health Database[71]	Personal health information from the Andalusian Population Health Database and Healthcare information	Health and healthcare information	Chronic diseases, functional and cognitive assessments, health resources (volume and cost), population stratification, drugs consumed
REDIAM - Andalusian Environmental Information Network[72]	Daily averages by collecting/meteorological station and at the census section level	Pollution, temperature	Mean daily values from pollution, air quality and temperature
SVEA - Andalusian Epidemiological Surveillance System[73]	Functional organization for health surveillance that collects, among other things, epidemiological information related to SARS-COV-2 infection	Epidemiological information of COVID-19	PCR result, symptoms date, close contact, healthcare professional, hospitalization unit (specifying ICU), date of admission and discharge, need of mechanical ventilation and clinical data

For further details, see Supplementary Material 3

Table 6. Health Care and Social Survey (ESSOC): Information blocks and variables entered in each measurement.

Subject area	1st measurement (M1)	2nd measurement (M2)	3rd measurement (M3)	4th and 5th measurements (M4 and M5)
Household and housing characteristics	Municipality, usual household, type of household, surface area, facilities, household changes, number of cohabitants (<6/<16/>60), type of household, and equipment.	Municipality, usual household, type of household, surface area, facilities, household changes[b], number of cohabitants (<6/<16/>60), equipment, number of rooms, and number of inhabitants with disabilities or requiring care.	Municipality, usual household, type of household, surface area, household changes, number of cohabitants (<6/<16/>60), number of rooms, and number of inhabitants with disabilities or requiring care.	Municipality, usual household, type of household, surface area, household changes, number of cohabitants (<6/<16/>60), number of rooms, and number of inhabitants with disabilities or requiring care.
Time use and cohabitation	Household chores, care tasks, daily activities during the confinement period (at home and outside), cohabitation and relationships, and causes for optimism.			Household chores, care tasks, daily activities[b] during the confinement period (at home and outside), cohabitation and relationships, and causes for optimism.
Health and emotional well-being	COVID-19 diagnosis, severity, diagnosis within the person's settings, self-perception of general and mental health (current and last year), emotional well-being[c], difficulty to withstand the confinement, malaise, chronic illness, and change of medication.	COVID-19 diagnosis, severity, diagnostic tests, diagnosis within the person's settings, self-perception of general and mental health, emotional well-being[b], cohabitation, difficulty to withstand the confinement, happiness, social and emotional support[c], malaise[b], chronic diseases (suffering and limitations), and medication (use and change of use[b]).	COVID-19 diagnosis[b], severity, diagnostic tests, diagnosis within the person's settings, self-perception of general and mental health, emotional well-being[bc], happiness, social and emotional support[bc], malaise[b], chronic diseases (suffering and limitations), and medication (use and change of use[b]).	COVID-19 diagnosis[b], severity, diagnostic tests, diagnosis within the person's settings, self-perception of general and mental health, emotional well-being[bc], happiness, social and emotional support[bc], malaise[b], chronic diseases (suffering and limitations), and medication (use and change of use[b]).
Habits and lifestyle	Habit modification (exercising, smoking, alcohol consumption, sleep, and diet).	Habit modification[b]: exercising, drinking, smoking, sleep, food, daily intake of vegetables and fruit, exercising, weight and height[c], smoking, alcohol consumption sleep, and flu vaccination.	Habit modification[b]: exercising, drinking, smoking, sleep, food, daily intake of vegetables and fruit, exercising, weight and height[c], smoking, alcohol consumption sleep, and flu vaccination[b]	Habit modification[b]: exercising, drinking, smoking, sleep, food, daily intake of vegetables and fruit, exercising, weight and height[c], smoking, alcohol consumption sleep, and flu and COVID-19 vaccination[b]
Economic situation and socio-demographic characteristics	Educational level, employment situation, working from home, type of contract, occupation[c], cohabitation with a partner, identification of the cohabitant with the greater income (educational level, employment situation, type of contract, occupation), difficulty in making ends meet, late payments, income, future worries, and degree of confidence in public institutions.	Employment situation, educational level, occupation[c], development[b], ability to work, identification of the cohabitant with the greater income (educational level, employment situation, occupation), difficulty in making ends meet, late payments, change in economic situation, parents' educational level, and future worries.	Employment situation, educational level, occupation[c], development, ability to work, identification of the cohabitant with the greater income (educational level, employment situation, occupation), difficulty in making ends meet, late payments[b], change in economic situation, parents' educational level, and future worries.	Employment situation, educational level, occupation[c], development, ability to work, identification of the cohabitant with the greater income (educational level, employment situation, occupation), difficulty in making ends meet, late payments[b], change in economic situation, parents' educational level, and future worries.

[a] Questionnaires are provided as Supplementary Material 2
[b] Variables that present modifications in their temporal scope in relation to the previous measurement.
[c] Composite variables: emotional well-being[75], social and emotional support[76,77], body mass index [78], social class[79,80]

2.11. Data Analysis

The analyses take advantage of all the information available from the measurements and the auxiliary information sources and will be carried out with the free software environment R[81], considering the sample design, as well as the calibration and inference methods described in the previous sections. The use of free software will guarantee transparency and facilitate the replicability of the study.

A table will be prepared for each variable of each measurement, together with the variable's original response categories, including the valid sample size (n), the percentage of lost samples, the population size estimate (N), the relevant statistic (mean or percentage), the 95% CI, and the coefficient of variation (CV), for both the total and the cross-disaggregation per sex and age (16–29/30–44/45–64/65+), as well as per province and urbanization degree. The sample size is recorded for the total and the categories of the segmentation variable. In the case of cells with CV estimates >20%, the CV will be indicated in a footnote to the table.

The variables shared by all measurements are dichotomized based on the results reported in the previous tables, identifying, in each case, the most convenient category to be highlighted based on the previous tables. In addition, a table describing the specific estimates with their corresponding CV will also be created.

Alternatively, to evaluate the changes in each measurement with respect to the first one, both the population affected by the change and the percentage segmented per demographic and territorial variables will be estimated. The p value will be calculated to evaluate the effect of such change and will be indicated in a footnote to the table using three categories: $p < 0.001$, $p < 0.05$, and $p < 0.1$.

In the case of variables that coincide in each pair of consecutive measurements (M2–M1, M3–M2, and M4–M3), the estimated percentage of the difference between one measurement and the previous one will be calculated in addition to the estimate of the population size and the signalling when the CV is greater than 20% and segmented by the demographic and territorial variables.

To analyse factors associated with variables of a given measurement or variables measuring the change between one measurement and another, we will use multivariate explanatory models adapted to the characteristics and nature of the variables and specified as generalized linear mixed models (GLMM)

with a family dependent on the type of dependent variable: Gaussian, when the variable is continuous (equivalent to a linear regression); binomial, when the variable is dichotomous (equivalent to a logistic regression); or Poisson, when the variable is discrete (equivalent to a Poisson regression). Random effects will be included in these models to capture the effects of unobserved confounders. Inferences will be made following a Bayesian perspective and using the integrated nested Laplace approximation (INLA)[82,83]. We will use penalized complexity priors known as PC priors. These priors are robust in the sense that they do not impact the results, and, in addition, they allow for an epidemiological interpretation[84]. The analyses will be carried out using free software R (version 4.0.4 or greater)[81], through the INLA package[82,83,85].

Finally, advanced data visualizations will be used to allow an in-depth exploratory analysis of the evolution of the study variables and a representation of the main results of the produced models. These visualizations will be developed using Python[86] programming language and integrated into software and web solutions that allow for interaction and dissemination.

2.12. Data Management Plan

The data management plan is provided as Supplementary Material 1. The type and format of data that will be collected and generated within the scope of this project is described in this plan, together with the procedure provided to access data (by whom, how, and when it can be accessed), data ownership, repository to deposit data, and procedure planned to guarantee the specific ethic and legal requirements.

Details of the Data Protection Impact Assessment (DPIA) will also be presented in here in accordance with the specific adaptation of this methodology to research projects in the health care sector[87,88,89]. Thus, the need for a DPIA was confirmed from the outset (Table S1). Subsequently, the data lifecycle was defined (Table S2), and the need and proportionality of the processing were analysed (Table S3) and, finally, a risk assessment and action plan developed (Table S4).

2.13. Scoping Review

The final objective will be achieved through a Scoping Review, whereby the existing evidence on survey-based research related to the measurement of the impact COVID 19 (from the outset of the pandemic) has had on health and its determinants are summarized. The review will be carried out using free terms and controlled language in the databases Pubmed, Scopus, WoS, EMBASE, CINAHL, PsyInfo, LiLac, OpenGray, Gray Literature Report and, likewise, through a free search in Google and institutional websites to locate institutional documents, abstracts, conferences or in any other format where studies and research work can be found. In a first phase, the research work will be selected independently and blindly in pairs, and the study populations, sample size and main objective information will be identified in order to collect the methodological characteristics, including elements on epidemiological design, auxiliary sources of information, methodologies, and topics addressed. This review will be carried out at the beginning of the project and updated throughout its duration, which will allow the development of the rest of activities, especially the identification of hypotheses and the application of other methodologies to be guided, as well as define the lines of research and propose more appropriate methods for future research.

Finally, with the Open Science approach, the results, microdata, reports, codes, and documentation related to the Scoping Review will be accessible through an open web platform. Likewise, scientific articles will be sent to Open Access journals and organised events will be aimed at both the scientific community and the general population. There will be accessible dissemination of the data on the Internet through institutional websites and social-media channels and a virtual practice community will be created (from the Mendeley citation manager) to share references identified by the research team. The scientific community and anyone else who might be interested will have access to all of the above.

3. Discussion

Given the characteristics, current and future repercussions of the current pandemic, developing this research project will make it possible to periodically obtain relevant information for decision-making processes in social and health matters and, therefore, promote a more efficient, reliable, and responsible science for social change (like the one that we are currently experiencing as a result of the pandemic). In addition, it will encourage:

- institutional alliances of great social value between the public administration, health care services, and the scientific and academic community;
- powerful and novel methodologies in the fields of public health, epidemiology, and sampling to reduce potential biases caused by a lack of coverage, response, or non-randomized selection[1, 50];
- large, integrated, quality, and open databases containing information extracted from clinical and non-clinical population registries; data concerning social, economic, and environmental contexts and the perception of the population, along with foreseeing the future incorporation of genomic information [90];
- the systematic review throughout the entire project of the scientific evidence obtained through this type of study;
- training with a view to transmitting the available knowledge and increasing capacities and skills in designs, sources, and methodologies; and
- measuring the short- and mid-term impact of COVID-19 at different times and on different populations since the beginning of the official State of Alarm.

The limitations of the study include those derived from the coverage and quality of the sampling frames, which may cause selection biases. In the case of the BDLPA, its telephone coverage is over 90% and it tends to have low percentages of non-existent telephone numbers (7%–8%) or non-contactable

telephone numbers due to the frame's outdatedness (9%–10%). That said, any potential biases due to such defects will be corrected through the estimator calibration techniques described above. Another limitation is that caused by the longitudinal design of the ESSOC surveys in terms of panel attrition, which could lead to potential biases due to an absence of response. We estimate the response rate for the full longitudinal sample to be 25%–35% (the effective sample in M5 of the effective sample in M1 is divided by the effective sample in M1). To reduce the effect of potential non-response biases, we decided, on the one hand, to use an overlapping panel design as this type of design allows for completing each measurement with new cross-sectional samples that become longitudinal in subsequent measurements and, on the other hand, to make adjustments in each measurement according to the weights of the longitudinal samples.

In addition, this project may have information biases intrinsic to sampling research. Consequently, we have chosen to employ scales and variables widely used in population-based health surveys and that also allow comparisons to be made for most of the key indicators. Finally, as in most of these types of studies, its feasibility and limitations will, to a large extent, depend on the quality and availability of the data from different data holders. Data availability for this project would appear not to be a limitation, given that the request for the processing of these data is in accordance with the policies on data disclosure for processing data in research projects in the public sector. Additionally, the EIPD performed resulted in an acceptable residual risk level (see Supplementary Material 1). This evaluation highlights, among other elements, that the data provided is protected by statistical confidentiality, that it is not misused, that its treatment is anonymous and global at all times, and that indirect identification is impossible. Furthermore, the project results will be beneficial to the general population in a holistic way, thanks to its socioeconomic and environmental context, and its evolution over several years from the onset of the pandemic. Therefore, the risk to the privacy of the study population is minimal compared to the potential benefit of its results. The transfer of data will be carried out between organizations within the Junta de Andalucía, as this will be within the context of a research project exclusively in the public sphere and under the legitimate use of records as research infrastructures in accordance with the Spanish health legislation (see Supplementary Material 1). With respect to data quality, the reproducibility of this research will be conditioned by the way in which the data have been obtained, a phenomenon common to all RWD studies. Even when these are the best possible, the fact that the data are extracted from official sources does not guarantee their quality. Because of this, the owners of the information will be asked to describe the detailed procedure used to extract the data, as well as any previous processing it may have undergone. It should be noted that individuals collaborating on this project also belong to the work teams of the main sources of information to be used which, therefore, should guarantee success in the interpretation of the original data and the optimization of the extraction strategies.

4. Conclusions

The ESSOC will enable precise and valid analysis of the short- and mid-term impact of the policies applied, and interventions made to, not only of the health of the general population, but also the most vulnerable population, during the pandemic. The study will also determine the evolution of the socioeconomic, psychosocial, behavioural, occupational, environmental, and clinical determinants of health and identify the inequalities in health in all its axes (i.e., social class, gender, age, ethnicity, and territory).

The conceptual approach of this study will encompass all aspects affecting health and so will contribute to an extraordinary increase in the current knowledge concerning the impact the COVID-19 pandemic is having. This knowledge will, in turn, be crucial for health systems in order to design quick and effective interventions aimed at improving the health care, health, and quality of life of the populations most affected by the COVID-19 pandemic. Moreover, the project management model based on collaborative, multidisciplinary, and open research will allow the critical mass needed to be generated to thus achieve the objectives that have been set (i.e., at the populational level, as well as at an individual and disease level, in our case, COVID-19).

5. Declarations

Supplementary Materials: The Data Management Plan is available online at www.mdpi.com/xxx/s1. Questionnaires are available online at www.mdpi.com/xxx/s2 (Spanish version). Variables from administrative registries are available online at www.mdpi.com/xxx/s3 (Spanish version).

Author Contributions: conceptualization, A.C.L.; methodology, A.C.L, C.S.C, M.R, R.F, L.C., M.S. and I.E.; resources, I.E., M.F., R.V., N.L.; data curation, C.S.C., R.F. and L.C.; writing—original draft preparation, A.C.L. and C.S.C.; writing—review and editing, M.R, I.E., M.S., M.F., N.L, M.B., R.F., L.C., A.D.C. and R.V.; project administration, A.C.L. and C.S.C; funding acquisition, A.C.L. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: The funding received to date for this research is partly public, obtained from funds of the Andalusian Institute of Statistics and Cartography (IECA) and the Andalusian School of Public Health (EASP), and partly private, obtained from the competitive calls of the SUPERA COVID19 Fund of Santander Universities (SAUN), the Conference of Spanish University Rectors (CRUE, Conferencia de Rectores de Universidades Españolas), and the Spanish National Research Council (CSIC, Consejo Superior de Investigaciones Científicas), in addition to the COVID-19 Competitive Grant Program from Pfizer Global Medical Grants. The private funding sources did not participate in the design or conduct of the study; the collection, management, analysis, or interpretation of the data; or the preparation, review, or approval of the manuscript.

Acknowledgments: This study was carried out within the 'Cohort Real-World Data' subprogram of the CIBER of Epidemiology and Public Health (CIBERESP). It also forms part of a PhD thesis (C.S.C) conducted in the Health Sciences Doctoral Studies program of the Universities of Seville and Jaén (Spain). We would like to thank Eva Galindo from Demométrica for her support in the fieldwork, and Juan Carlos Fernández Merino, Cristina Torro García-Morato, Mª Dolores Muñoyerro and Javier García León from the Consejería de Salud y Familias de la Junta de Andalucía for their institutional support.

Institutional Review Board Statement: The study was approved by the Research Ethics Committee of the Department of Health and Families of the Andalusian Regional Government (protocol code 10/20, dated 07 December 2020). The ESSOC is an activity included in the processing activities registry of the Department of Health and Families of the Andalusian Regional Government and is linked to the Andalusian Health Survey (EAS, Encuesta Andaluza de Salud), an official statistical operation included in the Andalusian Statistical and Cartographic Plan (Plan Estadístico y Cartográfico de Andalucía), with code 02.02.21.

Informed Consent Statement: An informed consent was obtained from all subjects involved in the study. Further information can be found in the Data Management Plan (Supplementary Material 1).

Data Availability Statement: The data, source codes, and other documentation developed in the context of this study will be available through the web platform http://researchprojects.es/ and from the open source project management tool https://osf.io/.

Conflicts of Interest: The authors declare no conflict of interest and the funders had no role in the design of the study; the collection, analyses, or interpretation of the data; the writing of the manuscript; or in the decision to publish its results.

Abbreviations

The following abbreviations are used in this manuscript:

BDLPA: Longitudinal Andalusian Population Database

BDU: User Database from the Andalusian Health Survey

BPS: Andalusian Population Health Database

CIBERESP: Network Biomedical Research Centre of Epidemiology and Public Health

CSyF: Department of Health and Families of the Andalusian Regional Government

DPIA: Data Protection Impact Assessment

EASP: Andalusian School of Public Health

ESSOC: Health Care and Social Survey

IECA: Andalusian Institute of Statistics and Cartography

MNP: Natural Movement of the Population

REDIAM: Andalusian Environmental Information Network

RWD: Real-World Data

SVEA: Andalusian Epidemiological Surveillance System

References

1. Department of Social Policy and Intervention. University of Oxford. Surveys | Oxford Supertracker. The Global Directory for COVID Policy Trackers and Surveys [Internet]. [cited 2021 Mar 25]. Available from: https://supertracker.spi.ox.ac.uk/surveys/

2. MRC C for GIDA. COVID-19 reports | Faculty of Medicine | Imperial College London [Internet]. [cited 2021 Mar 14]. Available from: http://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/

3. ISCIII. EpiGraph – COVID-19 simulator [Internet]. [cited 2021 Mar 14]. Available from: https://www.arcos.inf.uc3m.es/epigraph/

4. IECA. Portal IECA sobre el COVID19 en Andalucía. Datos Sanitarios | Instituto de Estadística y Cartografía de Andalucía [Internet]. [cited 2021 Mar 14]. Available from: http://www.juntadeandalucia.es/institutodeestadisticaycartografia/salud/datosSanitarios.html

5. Prieto-Alhambra D, Balló E, Coma E, Mora N, Aragón M, Prats-Uribe A, et al. Hospitalization and 30-day fatality in 121,263 COVID-19 outpatient cases [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Mar 14]. p. 2020.05.04.20090050. Available from: https://doi.org/10.1101/2020.05.04.20090050

6. Burn E, You SC, Sena AG, Kostka K, Abedtash H, Abrahão MTF, et al. Deep phenotyping of 34,128 patients hospitalised with COVID-19 and a comparison with 81,596 influenza patients in America, Europe and Asia: An international network study [Internet]. Vol. 21, medRxiv. medRxiv; 2020 [cited 2021 Mar 14]. p. 26. Available from: https://doi.org/10.1101/2020.04.22.20074336

7. Xiong J, Lipsitz O, Nasri F, Lui LMW, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. Vol. 277, Journal of Affective Disorders. Elsevier B.V.; 2020. p. 55–64.

8. Orte C, Sánchez-Prieto L, Domínguez DC, Barrientos-Báez A. Evaluation of Distress and Risk Perception Associated with COVID-19 in Vulnerable Groups. Int J Environ Res Public Health [Internet]. 2020 Dec 9 [cited 2021 Mar 14];17(24):9207. Available from: https://www.mdpi.com/1660-4601/17/24/9207

9. Rose TC, Mason K, Pennington A, McHale P, Buchan I, Taylor-Robinson DC, et al. Inequalities in COVID19 mortality related to ethnicity and socioeconomic deprivation [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Mar 14]. p. 2020.04.25.20079491. Available from: https://doi.org/10.1101/2020.04.25.20079491

10. González-Sanguino C, Ausín B, Castellanos MÁ, Saiz J, López-Gómez A, Ugidos C, et al. Mental health consequences during the initial stage of the 2020 Coronavirus pandemic (COVID-19) in Spain. Brain Behav Immun [Internet]. 2020 Jul 1 [cited 2020 Nov 6];87:172–6. Available from: /pmc/articles/PMC7219372/?report=abstract

11. Ministerio de la Presidencia R con las C y MD. Real Decreto 463/2020 [Internet]. BOE. 2020 [cited 2021 Mar 3]. Available from: https://boe.es/buscar/pdf/2020/BOE-A-2020-3692-consolidado.pdf

12. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Vol. 395, The Lancet. Lancet Publishing Group; 2020. p. 912–20.

13. Jeong H, Yim HW, Song YJ, Ki M, Min JA, Cho J, et al. Mental health status of people isolated due to Middle East Respiratory Syndrome. Epidemiol Health [Internet]. 2016 [cited 2021 Mar 14];38:e2016048. Available from: /pmc/articles/PMC5177805/

14. Ricci-Cabello I, Meneses-Echavez JF, Serrano-Ripoll MJ, Fraile-Navarro D, de Roque MAF, Moreno GP, et al. Impact of viral epidemic outbreaks on mental health of healthcare workers: A rapid systematic review [Internet]. Vol. 4, medRxiv. medRxiv; 2020 [cited 2021 Mar 3]. p. 2020.04.02.20048892. Available from: https://doi.org/10.1101/2020.04.02.20048892

15. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly [Internet]. Vol. 5, The Lancet Public Health. Elsevier Ltd; 2020 [cited 2021 Mar 14]. p. e256. Available from: www.thelancet.com/public-health

16. Luo M, Guo L, Yu M, Wang H. The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public – A systematic review and meta-analysis [Internet]. Vol. 291, Psychiatry Research. Elsevier Ireland Ltd; 2020 [cited 2021 Mar 14]. p. 113190. Available from: /pmc/articles/PMC7276119/

17. Pappa S, Ntella V, Giannakas T, Giannakoulis VG, Papoutsi E, Katsaounou P. Prevalence of depression, anxiety, and insomnia among healthcare workers during the COVID-19 pandemic: A systematic review and meta-analysis [Internet]. Vol. 88, Brain, Behavior, and Immunity. Academic Press Inc.; 2020 [cited 2021 Mar 14]. p. 901–7. Available from: /pmc/articles/PMC7206431/

18. Drosten C. Entrevista Christian Drosten. eldiario.es [Internet]. 2020 [cited 2021 Mar 14]; Available from: https://www.eldiario.es/internacional/theguardian/christian-drosten-principal-coronavirus-hundiendo_128_5948822.html

19. Fundación Secretariado Gitano. Encuesta impacto social de la COVID-19 [Internet]. 2020 [cited 2021 Mar 14]. Available from: https://www.gitanos.org/actualidad/archivo/131067.html

20. Federación de Asociaciones de Mujeres Gitanas. Informe antigitanismo y COVID 19. [Internet]. [cited 2021 Mar 14]. Available from: https://fakali.org/wp-content/uploads/2020/12/Informe-FAKALI-Antigitanismo-y-Covid-19.pdf

21. Vallejo-Slocker L, Fresneda J, Vallejo MA. Psychological wellbeing of vulnerable children during the COVID-19 pandemic. Psicothema [Internet]. 2020 [cited 2021 Mar 14];32(4):501–7. Available from: http://www.psicothema.com/pdf/4628.pdf

22. Prats-Uribe A, Paredes R, Prieto-Alhambra D. Ethnicity, comorbidity, socioeconomic status, and their associations with COVID-19 infection in England: A cohort analysis of UK Biobank data [Internet]. medRxiv. medRxiv; 2020 [cited 2021 Mar 14]. p. 2020.05.06.20092676. Available from: https://doi.org/10.1101/2020.05.06.20092676

23. Guijarro C, Pérez-Fernández E, González-Piñeiro B, Meléndez V, Goyanes MJ, Renilla ME, et al. Differential risk for COVID-19 in the first wave of the disease among Spaniards and migrants from different areas of the world living in Spain. Rev Clin Esp. 2020 Nov 20;

24. Platt L, Warwick R. Are some ethnic groups more vulnerable to COVID-19 than others? Inequality: the IFS Deaton Review [Internet]. 2020 [cited 2021 Apr 11]; Available from: https://www.ifs.org.uk/inequality/chapter/are-some-ethnic-groups-more-vulnerable-to-covid-19-than-others/

25. Ruiz-Azarola A, Escudero Carretero M, López-Fernández LA, Gil García E, March Cerdà JC, López Jaramillo D. The perspective of migrants on access to health care in the context of austerity policies in Andalusia (Spain). Gac Sanit [Internet]. 2020 May 1 [cited 2021 Mar 14];34(3):261–7. Available from: https://doi.org/10.1016/j.gaceta.2018.09.006

26. Holt E. COVID-19 lockdown of Roma settlements in Slovakia. Lancet Infect Dis [Internet]. 2020 Jun 1 [cited 2021 Apr 11];20(6):659. Available from: www.thelancet.com/infection

27. Conde Blanco E, Manzanares I, Centeno M, Khawaja M, Betrán O, Donaire A, et al. Epilepsy and lockdown: A survey of patients normally attending a Spanish centre [Internet]. Vol. 143, Acta Neurologica Scandinavica. Blackwell Publishing Ltd; 2021 [cited 2021 Mar 14]. p. 206–9. Available from: https://pubmed.ncbi.nlm.nih.gov/32990951/

28. López-Bravo A, García-Azorín D, Belvís R, González-Oria C, Latorre G, Santos-Lasaosa S, et al. Impact of the COVID-19 pandemic on headache management in Spain: an analysis of the current situation and future perspectives. Neurologia [Internet]. 2020 Jul 1 [cited 2021 Apr 11];35(6):372–80. Available from: https://doi.org/10.1016/j.nrl.2020.05.006

29. Lara B, Carnes A, Dakterzada F, Benitez I, Piñol-Ripoll G. Neuropsychiatric symptoms and quality of life in Spanish patients with Alzheimer's disease during the COVID-19 lockdown. Eur J Neurol [Internet]. 2020 Sep 24 [cited 2021 Mar 14];27(9):1744–7. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ene.14339

30. Elbarbary NS, dos Santos TJ, de Beaufort C, Agwu JC, Calliari LE, Scaramuzza AE. COVID-19 outbreak and pediatric diabetes: Perceptions of health care professionals worldwide. Pediatr Diabetes [Internet]. 2020 Nov 1 [cited 2021 Mar 14];21(7):1083–92. Available from: https://pubmed.ncbi.nlm.nih.gov/32686287/

31. Pleguezuelos E, Del Carmen A, Moreno E, Ortega P, Vila X, Ovejero L, et al. The experience of COPD patients in lockdown due to the COVID-19 pandemic. Int J COPD [Internet]. 2020 [cited 2021 Mar 14];15:2621–7. Available from: /pmc/articles/PMC7591044/

32. Beisani M, Vilallonga R, Petrola C, Acosta A, Casimiro Pérez JA, García Ruiz de Gordejuela A, et al. Effects of COVID-19 lockdown on a bariatric surgery waiting list cohort and its influence in surgical risk perception. Langenbeck's Arch Surg [Internet]. 2020 Mar 1 [cited 2021 Mar 14];1. Available from: /pmc/articles/PMC7690848/

33. Hartmann-Boyce J, Mahtani KR. Supporting people with long-term conditions (LTCs) during national emergencies - The Centre for Evidence-Based Medicine [Internet]. [cited 2021 Mar 14]. Available from: https://www.cebm.net/covid-19/supporting-people-with-long-term-conditions-ltcs-during-national-emergencies/

34. Prieto M, March J, Martín A, Escudero M, López M, Luque N. Repercusiones del confinamiento por COVID-19 en pacientes crónicos de Andalucía [Internet]. Comprender el COVID-19 desde una perspectiva de Salud Pública. Escuela Andaluza de Salud Pública. 2020 [cited 2021 Apr 11]. Available from: https://www.easp.es/web/coronavirusysaludpublica/pacientes-cronicos-en-casa-la-experiencia-del-confinamiento-como-se-cuidan-que-necesitan-como-apoyarles/

35. Cabrera-León A, Sánchez-Cantalejo C. Características y resultados de encuestas sobre el impacto de la enfermedad COVID-19 | Comprender el COVID-19 desde una perspectiva de salud pública [Internet]. 2020 [cited 2021 Mar 15]. Available from: https://www.easp.es/web/coronavirusysaludpublica/caracteristicas-y-resultados-de-encuestas-sobre-el-impacto-de-la-enfermedad-covid-19/

36. Mansilla Domínguez JM, Font Jiménez I, Belzunegui Eraso A, Peña Otero D, Díaz Pérez D, Recio Vivas AM. Risk Perception of COVID–19 Community Transmission among the Spanish Population. Int J Environ Res Public Health [Internet]. 2020 Dec 2 [cited 2021 Apr 11];17(23):8967. Available from: https://www.mdpi.com/1660-4601/17/23/8967

37. Kim JK, Wang Z. Sampling Techniques for Big Data Analysis. Int Stat Rev [Internet]. 2019 May 31 [cited 2021 Mar 15];87(S1):S177–91. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12290

38. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet [Internet]. 2020 Aug 22 [cited 2020 Nov 6];396(10250):535–44. Available from: https://pubmed.ncbi.nlm.nih.gov/32645347/

39. Pastor-Barriuso R, Pérez-Gómez B, Hernán MA, Pérez-Olmeda M, Yotti R, Oteo-Iglesias J, et al. Infection fatality risk for SARS-CoV-2 in community dwelling population of Spain: Nationwide seroepidemiological study. BMJ [Internet]. 2020 Nov 27 [cited 2021 Mar 15];371. Available from: https://www.bmj.com/content/371/bmj.m4509

40. ISCIII. ESTUDIO ENE-COVID: INFORME FINAL ESTUDIO NACIONAL DE SERO-EPIDEMIOLOGÍA DE LA INFECCIÓN POR SARS-COV-2 EN ESPAÑA [Internet]. [cited 2021 Mar 15]. Available from: http://www.thelancet.com/journals/lancet/article/PIIS0140-6736

41. Ministerio de Sanidad. Centro de Coordinación de Alertas y Emergencias Sanitarias. Actualización nº 96. Enfermedad por el coronavirus (COVID-19). 05.05.2020 (datos consolidados a las 21:00 horas del 04.05.2020). SITUACIÓN EN ESPAÑA. [Internet]. [cited 2021 Mar 15]. Available from: https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Actualizacion_96_COVID-19.pdf

42. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis [Internet]. 2020 Jun 1 [cited 2021 Mar 15];20(6):669–77. Available from: www.thelancet.com/infectionVol

43. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 — COVID-NET, 14 States, March 1–30, 2020. MMWR Morb Mortal Wkly Rep [Internet]. 2020 Apr 17 [cited 2021 Mar 15];69(15):458–64. Available from: http://www.cdc.gov/mmwr/volumes/69/wr/mm6915e3.htm?s_cid=mm6915e3_w

44. Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, et al. Prevalence of comorbidities and its effects in coronavirus disease 2019 patients: A systematic review and meta-analysis. Int J Infect Dis [Internet]. 2020 May 1 [cited 2021 Mar 15];94:91–5. Available from: https://pubmed.ncbi.nlm.nih.gov/32173574/

45. Vardavas CI, Nikitara K. COVID-19 and smoking: A systematic review of the evidence [Internet]. Vol. 18, Tobacco Induced Diseases. International Society for the Prevention of Tobacco Induced Diseases; 2020 [cited 2021 Mar 15]. Available from: /pmc/articles/PMC7083240/

46. Pinazo-Hernandis S. Psychosocial impact of COVID-19 on older people: Problems and challenges. Rev Esp Geriatr Gerontol [Internet]. 2020 Sep 1 [cited 2021 Mar 15];55(5):249–52. Available from: /pmc/articles/PMC7266768/

47. Losada-Baltar A, Martínez-Huertas J, Jiménez-Gonzalo L, Pedroso-Chaparro M, Gallego-Alberto L, Fernandes-Pires J, et al. Longitudinal Correlates of Loneliness and Psychological Distress During the Lockdown Situation due to COVID-19. Effects of Age and Self-Perceptions of Aging. Journals Gerontol Ser B Psychol Sci Soc Sci [Internet]. 2021 [cited 2021 Mar 15]; Available from: https://pubmed.ncbi.nlm.nih.gov/33438002/

48. Arpino B, Pasqualini M, Bordone V, Solé-Auró A. Older People's Nonphysical Contacts and Depression During the COVID-19 Lockdown. G Castle N, editor. Gerontologist [Internet]. 2021 Feb 23 [cited 2021 Apr 11];61(2):176–86. Available from: https://academic.oup.com/gerontologist/article/61/2/176/5911810

49. Durán JC et al. Atención al mayor, abordaje multidisciplinar de su complejidad. (SAGG) SA de G y G, editor. 2019.

50. Cabrera-León A, Daponte Codina A, Mateo I, Arroyo-Borrell E, Bartoll X, Bravo MJ, et al. Indicadores contextuales para evaluar los determinantes sociales de la salud y la crisis económica española. Gac Sanit. 2017 May 1;31(3):194–203.

51. Hernán M, Robins J. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC [Internet]. Chapman & Hall/CRC, editor. 2020. Available from: https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2021/01/ciwhatif_hernanrobins_31jan21.pdf

52. Beaumont J. Are probability surveys bound to disappear for the production of official statistics? [Internet]. Statistics Canada. 2020 [cited 2021 Apr 11]. Available from: https://www150.statcan.gc.ca/n1/pub/12-001-x/2020001/article/00001-eng.htm

53. Burgelman J-C, Pascu C, Szkuta K, Von Schomberg R, Karalopoulos A, Repanas K, et al. Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. Front Big Data [Internet]. 2019 Dec 10 [cited 2021 Mar 15];2:43. Available from: https://www.frontiersin.org/article/10.3389/fdata.2019.00043/full

54. Kalton G, Citro CF. Panel Surveys: Adding the Fourth Dimension. Innov Eur J Soc Sci Res [Internet]. 1995 [cited 2021 Mar 25];8(1):25–39. Available from: https://www.tandfonline.com/doi/abs/10.1080/13511610.1995.9968429

55. Junta de Andalucía. Junta de Andalucía - Zonas desfavorecidas [Internet]. [cited 2021 Apr 30]. Available from: https://www.juntadeandalucia.es/organismos/igualdadpoliticassocialesyconciliacion/areas/inclusion/zonas-transformacion.html

56. Merlo J, Viciana-Fernández FJ, Ramiro-Fariñas D. Bringing the individual back to small-area variation studies: A multilevel analysis of all-cause mortality in Andalusia, Spain. Soc Sci Med. 2012 Oct;75(8):1477–87.

57. Ieca I de E y C de A. Base de Datos Longitudinal de Población de Andalucía [Internet]. 2018. Available from: https://www.juntadeandalucia.es/institutodeestadisticaycartografia/metodologias/IME010101.pdf LB - AyKB

58. Junta de Andalucía C de S y F. Base de Datos de Usuarios del Sistema Sanitario Público de Andalucía [Internet]. 2018. Available from: https://www.juntadeandalucia.es/organismos/saludyfamilias/servicios/estadistica-cartografia/actividad/detalle/175259.html LB - AxE8

59. Consorcio Fernando de los Ríos. Guadalinfo [Internet]. [cited 2021 Mar 15]. Available from: http://www.guadalinfo.es/

60. Junta de Andalucía. ESTRATEGIA REGIONAL ANDALUZA PARA LA COHESIÓN E INCLUSIÓN SOCIAL.

61. IECA. Clasificación del Grado de Urbanización | Instituto de Estadística y Cartografía de Andalucía [Internet]. [cited 2021 Mar 15]. Available from: http://www.juntadeandalucia.es/institutodeestadisticaycartografia/gradourbanizacion/index.htm

62. Singh S. Advanced sampling theory with applications. Springer, editor. Science & Business Media; 2003.

63. Yves Tillé A, Matei A, Alina Matei M. Package "sampling" Title Survey Sampling. 2021.

64. INE. La encuesta continua de hogares [Internet]. 2019. Available from: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176952&menu=ultiDatos&idp=1254735572981 LB - sijq

65. Plewis I. Non-response in a birth cohort study: The case of the millennium cohort study. Int J Soc Res Methodol [Internet]. 2007 Dec [cited 2021 Mar 15];10(5):325–34. Available from: https://www.tandfonline.com/doi/abs/10.1080/13645570701676955

66. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. New York, NY, USA: Association for Computing Machinery; 2016 [cited 2021 Apr 11]. p. 785–94. Available from:

https://dl.acm.org/doi/10.1145/2939672.2939785

67. Kott PS, Liao D. One step or two? Calibration weighting from a complete list frame with nonresponse. Stat Canada [Internet]. 2015 [cited 2021 Mar 15];41(1):165–81. Available from: www.statcan.gc.ca

68. Kott PS, Liao D. Calibration Weighting for Nonresponse that is Not Missing at Random: Allowing More Calibration than Response-Model Variables. J Surv Stat Methodol [Internet]. 2017 Jun 1 [cited 2021 Mar 15];5(2):159–74. Available from: https://academic.oup.com/jssam/article-lookup/doi/10.1093/jssam/smx003

69. Ferri-García R, Del Mar Rueda M. Efficiency of propensity score adjustment and calibration on the estimation from non-probabilistic online surveys. SORT [Internet]. 2018 Jul 1 [cited 2021 Mar 15];42(2):159–82. Available from: https://ddd.uab.cat/record/200758

70. Ferri-García R, Rueda M del M. Propensity score adjustment using machine learning classification algorithms to control selection bias in online surveys. Marchetti S, editor. PLoS One [Internet]. 2020 Apr 22 [cited 2021 Mar 15];15(4):e0231500. Available from: https://dx.plos.org/10.1371/journal.pone.0231500

71. Servicio Andaluz de Salud. Base poblacional de salud | Servicio Andaluz de Salud [Internet]. [cited 2021 Mar 15]. Available from: https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/profesionales/sistemas-de-informacion/base-poblacional-de-salud

72. Consejeria De Medioambiente. Junta de Andalucía. REDIAM | IDEAndalucia [Internet]. [cited 2021 Mar 25]. Available from: https://www.ideandalucia.es/portal/nodo-rediam

73. Consejería de Salud y Familias. Junta de Andalucía. Programas de Vigilancia de Enfermedades Transmisibles [Internet]. [cited 2021 Mar 25]. Available from: https://www.juntadeandalucia.es/organismos/saludyfamilias/areas/salud-vida/vigilancia/paginas/vigilancia-transmisibles.html

74. Servicio Andaluz de Salud. Documento de identificacion de patologias crónicas de la BPS [Internet]. 2018 [cited 2021 Mar 15]. Available from: https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-mediafile_sasdocumento/2019/patologias_bps.pdf

75. Bericat E. EXCLUIDOS DE LA FELICIDAD. Las desigualdades de bienestar emocional en España y Europa. [Internet]. 2015 [cited 2021 Mar 15]. Available from: https://www.researchgate.net/publication/287814426_EXCLUIDOS_DE_LA_FELICIDAD_Las_desigualdades_de_bienestar_emocional_en_Espana_y_Europa

76. Bellón J, Delgado A, Luna J, Lardelli P. Validez y fiabilidad del cuestionario de apoyo social funcional Duke-UNC-11. Aten Primaria [Internet]. 1996 [cited 2021 Mar 15];18:153–63. Available from: https://www.elsevier.es/es-revista-atencion-primaria-27-articulo-validez-fiabilidad-del-cuestionario-apoyo-14325

77. Broadhead WE, Gehlbach SH, de Gruy F V., Kaplan BH. The Duke-UNC functional social support questionnaire: Measurement of social support in family medicine patients. Med Care [Internet]. 1988 [cited 2021 Mar 15];26(7):707–21. Available from: https://pubmed.ncbi.nlm.nih.gov/3393031/

78. OMS. 10 datos sobre la obesidad [Internet]. [cited 2021 Mar 15]. Available from: https://www.who.int/features/factfiles/obesity/facts/es/

79. Domingo-Salvany A, Bacigalupe A, Carrasco JM, Espelt A, Ferrando J, Borrell C. Propuestas de clase social neoweberiana y neomarxista a partir de la Clasificación Nacional de Ocupaciones 2011. Gac Sanit. 2013 May 1;27(3):263–72.

80. Chilet-Rosell E, Álvarez-Dardet C, Domingo-Salvany A. Utilización de las propuestas espã nolas de medición de la clase social en salud Use of Spanish proposals for measuring social class in health sciences. Gac Sanit. 2012;26(6):566–9.

81. R Core Team. R: The R Project for Statistical Computing [Internet]. 2021 [cited 2021 Mar 15]. Available from: https://www.r-project.org/

82. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B (Statistical Methodol [Internet]. 2009 Apr 1 [cited 2021 Mar 15];71(2):319–92. Available from: http://doi.wiley.com/10.1111/j.1467-9868.2008.00700.x

83. Rue H, Riebler A, Sørbye SH, Illian JB, Simpson DP, Lindgren FK. Bayesian Computing with INLA: A Review. Annu Rev Stat Its Appl [Internet]. 2017 Mar 7 [cited 2021 Mar 15];4(1):395–421. Available from: http://www.annualreviews.org/doi/10.1146/annurev-statistics-060116-054045

84. Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: A principled, practical approach to constructing priors. Stat Sci [Internet]. 2017 Feb 1 [cited 2021 Mar 15];32(1):1–28. Available from: http://arxiv.org/abs/1403.4630

85. R-INLA Project [Internet]. [cited 2021 Mar 15]. Available from: https://www.r-inla.org/home

86. Welcome to Python.org [Internet]. [cited 2021 Mar 15]. Available from: https://www.python.org/

87. Agencia Española de Protección de Datos. Guía práctica para LAS Evaluaciones de Impacto en la Protección de LOS datos sujetas al RGPD [Internet]. [cited 2021 Mar 15]. Available from: https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf

88. European Commission. Horizon 2020 Programme Guidance How to complete your ethics self-assessment [Internet]. 2019 [cited 2021 Mar 15]. Available from: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf

89. Agencia Española de Protección de Datos. Herramienta GESTIONA EIPD [Internet]. [cited 2021 Mar 15]. Available from: https://gestiona.aepd.es/

90. Consejería de Salud y Familias. Junta de Andalucía. Sistema Integrado de Epidemiología Genómica de Andalucía [Internet]. [cited 2021 Mar 26]. Available from: https://www.juntadeandalucia.es/organismos/saludyfamilias/areas/seguridad-alimentaria/salud-alimentos/paginas/siega.html

Figures

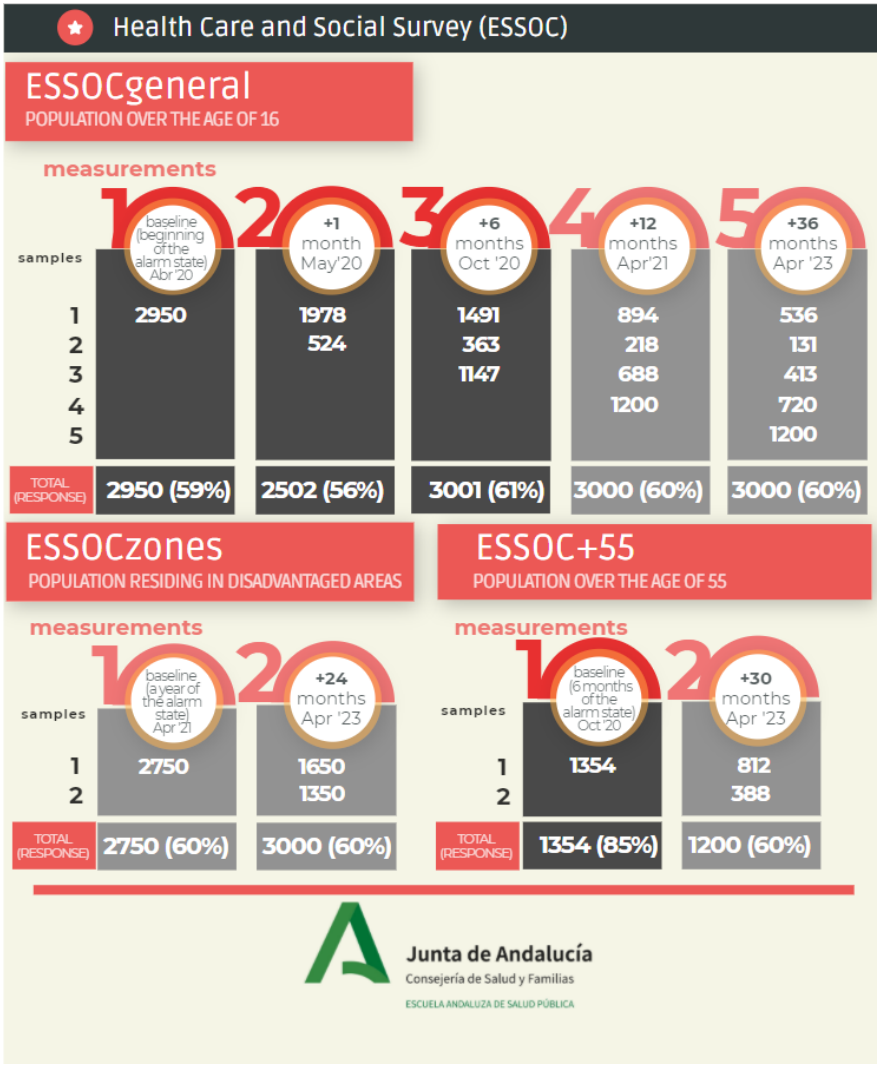

Figure 1

Health Care and Social Survey (ESSOC): Population and Temporal Scope, and Samples

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial1DataManagementPlan.pdf
- SupplementaryMaterial2Questionnaires.pdf
- SupplementaryMaterial3AuxiliaryVariables.pdf