



OPEN

Key factors affecting people's unwillingness to be confined during the COVID-19 pandemic in Spain: a large-scale population study

Marina Martinez-Garcia¹, Alejandro Rabasa², Xavier Barber², Kristina Polotskaya², Kristof Roomp³ & Nuria Oliver⁴✉

Population confinements have been one of the most widely adopted non-pharmaceutical interventions (NPIs) implemented by governments across the globe to help contain the spread of the SARS-CoV-2 virus. While confinement measures have been proven to be effective to reduce the number of infections, they entail significant economic and social costs. Thus, different policy makers and social groups have exhibited varying levels of acceptance of this type of measures. In this context, understanding the factors that determine the willingness of individuals to be confined during a pandemic is of paramount importance, particularly, to policy and decision-makers. In this paper, we study the factors that influence the unwillingness to be confined during the COVID-19 pandemic by the means of a large-scale, online population survey deployed in Spain. We perform two types of analyses (logistic regression and automatic pattern discovery) and consider socio-demographic, economic and psychological factors, together with the 14-day cumulative incidence per 100,000 inhabitants. Our analysis of 109,515 answers to the survey covers data spanning over a 5-month time period to shed light on the impact of the passage of time. We find evidence of *pandemic fatigue* as the percentage of those who report an unwillingness to be in confinement increases over time; we identify significant *gender* differences, with women being generally less likely than men to be able to sustain long-term confinement of at least 6 months; we uncover that the *psychological impact* was the most important factor to determine the willingness to be in confinement at the beginning of the pandemic, to be replaced by the *economic impact* as the most important variable towards the end of our period of study. Our results highlight the need to design gender and age specific public policies, to implement psychological and economic support programs and to address the evident pandemic fatigue as the success of potential future confinements will depend on the population's willingness to comply with them.

Population confinements have been one of the most widely adopted non-pharmaceutical interventions (NPIs) implemented by governments across the globe to help contain the spread of the SARS-CoV-2 virus¹⁻⁴. While confinement measures have been proven to be effective to reduce the number of infections^{3,5-7}, they entail significant economic and social costs. Thus, different policy makers and social groups have exhibited varying levels of acceptance of this type of measures, from the low levels of mask wearing compliance⁸ and public acceptance of COVID-19 lockdown scenarios⁹ in Germany to the highly interventionist Chinese experience^{10,11}. In this context, understanding the factors that determine the willingness of individuals to be confined during a pandemic is of paramount importance, particularly, to policy and decision-makers.

In this paper, we study the factors that determine people's unwillingness to be confined during the first five months of the COVID-19 pandemic in Spain by means of a large-scale, online population survey. We consider socio-demographic, economic and psychological factors, together with reported 14-day cumulative incidence

¹Dept de matemàtiques, Universitat Jaume I, Castelló 12071, Spain. ²Universidad Miguel Hernández, Centro de Investigación Operativa, Elche 03202, Spain. ³Microsoft Corporation, Redmond, WA, USA. ⁴ELLIS Unit Alicante Foundation, Alicante, Spain. ✉email: nuria@alum.mit.edu

per 100,000 inhabitants. Our analysis covers data spanning over a 5-month time period to shed light on the impact of the passage of time.

Since the COVID-19 pandemic was declared, numerous studies that have analysed the economic, social and psychological impact of NPIs. Moreover, several authors have reported on people's experiences during the confinements implemented to contain the spread of coronavirus. We summarize next the most relevant pieces of previous work as they relate to ours.

Economic impact of NPIs. Regarding the economic impact of non-pharmaceutical interventions, several works have assessed the cost of different types of NPIs in the context of the current coronavirus pandemic. Abiad et al. simulate different scenarios of the evolution of the pandemic in order to estimate the economic impact¹². Fernandes et al. maintain the same central idea of studying different scenarios and highlight how different this crisis is (due to the pandemic) compared to previous ones¹³. Possible supply chain scenarios are analyzed in¹⁴, highlighting that the timing of the closing and opening of the facilities at different stages might become a major factor in determining the impact of epidemic outbreaks on the supply chain performance. Finally, the work by Chen and colleagues¹⁵ highlights the impact that socio-psychological variables have on household consumption patterns and their ability and willingness to bear higher economic costs.

Social impact of NPIs. Beyond the economic dimension, the pandemic has also significantly impacted people's social lives. Clemens et al. assess the impact that social distance and confinement measures have on people's relationships¹⁶. Jiao and colleagues¹⁷ establish guidelines for children and adolescents to mitigate the effect of social isolation during the pandemic. Social vulnerability is pointed out as the main risk of the pandemic by¹⁸, recommending improvements of mental, social and spiritual resilience in the population. Mukhtar reviews past outbreaks to understand the impact of adverse effects on psychological health, psychological crisis intervention, and mental health management plans¹⁹. He concludes that individuals (and societies in general) must deliberately cultivate resilience and develop strategies to cope with adversity, as in the case of the current SARS-CoV-2 pandemic caused by confinement.

Psychological and behavioural impact of NPIs. Several studies have analyzed the psychological impact of the confinement measures via online surveys and/or smartphone apps. Gao et al.²⁰ report a significant increase in media usage during the first wave of the SARS-CoV-2 pandemic in China from a sample of 4,872 individuals collected over a 3-day period (from the 31st of January to the 2nd of February, 2020). Their main findings show that there is a high prevalence of mental health problems, positively associated with frequent exposure to social media during the COVID-19 outbreak. As a result of this finding, the authors suggest that governments should pay more attention to the population's mental health among the general population while fighting the coronavirus pandemic. Hawryluch et al.²¹ describe the results of analyzing data from 129 individuals collected via a long 152-question survey to capture the psychological impact of the confinement in Canada, during previous SARS-CoV-1 outbreak. They conclude that the prevalence of post-traumatic stress disorder and depressive symptoms grows as the period of confinement increases. A multivariate logistic regression model was used by Wang et al.²² to assess the impact of the confinement measures on people's mental health. The authors observed that the confinement measures adopted by the governments had a negative impact on the mental health of particularly vulnerable groups, including those with previous mental health disorders or chronic physical problems, first-line workers, coronavirus infected or suspected patients and those who had financial problems.

Recent work has also explored the use of technology to provide psychological support to people while in confinement, such as the self-guided activities proposed in²³, models to manage potential psychological crises and decision support systems to help alleviate psychiatric and psychological disorders caused or exacerbated by the confinement²⁴.

General behavioural patterns and perceptions towards the pandemic and the confinement measures are described in²⁵ from a sample of 1,420 individuals collected between the 18th and 24th of March, 2020 in Australia. The study identifies the main difficulties reported by the participants to comply with the measures of physical distancing, namely home sharing, the responsibility of caring for someone outside your home and working from home. Yamada et al. describe the COVIDiSTRESS survey (data collected between April and May 2020) deployed to study the psychological and behavioural responses to the coronavirus pandemic and associated government measures—such as the cancellation of public events and stay-at-home orders implemented in many countries. A global consortium of 39 countries and regions worked together to construct and translate the survey with variables of shared interests and recruited participants in 47 languages²⁶. The survey, with 1,304 respondents, covered several areas, including demographic data, physical symptoms in the past 14 days, contact history with COVID-19 in the past 14 days, knowledge and concerns about COVID-19, precautionary measures against COVID-19 in the past 14 days, additional information required with respect to COVID-19, the psychological impact of the COVID-19 outbreak; and mental health status. Finally, Bailon et al.²⁷ analyze the longitudinal answers of 1,000 respondents provided over a 2.5 month period in Spain, aiming to understand potential changes in people's mood during the confinement. Beyond the study on mood variations during confinement, this paper offers interesting descriptive statistics of the socio-demographic variables of the participants, as well as their housing characteristics and physical activity habits.

Willingness to comply with the confinement measures. Similar to our research, most of the previous work that has studied people's experience with the COVID-19 pandemic confinement measures has relied on self-reported survey data. Amat et al.²⁸ analyze 1,600 answers collected between the 20th and 28th of March, 2020 from Spain and United States. The authors claim that during the coronavirus pandemic individuals have

relinquished more rights than during other international crisis scenarios as climate change or international terrorism. Flexibility and the ability to adapt have been found to be key skills to help people cope with confinement situations. Tintori et al.²⁹ report such a finding from analysing 140,000 online interviews between the 22nd of March and the 2nd of April, 2020; and Landi et al.³⁰ collect data from 944 adults in Italy over 35 days and propose an acceptance and compromise therapy based in the concept of *psychological flexibility* as a tool to help people during confinement. McBride et al.³¹ propose empirical indicators captured by the Brief Resilience Scale to quantitatively assess people's acceptance of the confinement measures. They validate the proposed instrument via a study with 2,000 individuals in United States, between the 23rd and the 28th of March, 2020. Other authors, such as³², study the impact of conspiracy theories on people's willingness to comply with the confinement measures. They collect data via an online survey with 2,501 adults in England, which are sampled to match the distribution of English population by age, gender, income, and region. According to their findings, there is a notable endorsement of conspiracy beliefs about coronavirus in England, and such ideas do not appear to be limited to the fringes. These conspiracy beliefs connect to other forms of mistrust and are associated with less compliance with government guidelines and greater unwillingness to take up future tests and treatment. Finally, Chan et al.³³ suggest that regions with risk-averse attitudes are more likely to adjust their behavioural activity in response to the declaration of a pandemic even before official government lockdowns.

In view of the reported previous work and given the focus of our research, the main contributions of our paper are four-fold. First, we report the results of analyzing a very large-scale, online population survey with 109,515 answers from Spain collected between April 3rd, 2020 and September 11th, 2020. Second, we apply two different methodologies to predict and shed light on the factors that impact people's willingness to be confined: logistic regression and pattern-discovery, rule-based models. Third, we consider socio-demographic, economic and psychological factors, together with reported 14-day cumulative incidence per 100,000 inhabitants. Finally, we analyse data covering a 5-month period to understand the impact of the passage of time on people's unwillingness to be confined.

In the following section, we describe the main results of our analyses. Next, we provide a discussion of our main results, followed by the main conclusions derived from our work. We conclude with a description of the methodology that we followed.

Results

In this section, we describe the main results of our analyses, starting with the descriptive statistics of the data set, followed by the results of a logistic regression modeling and the patterns that were automatically identified by our pattern identification approach (described in the [Methods](#) section).

Descriptive statistics. We report the results of analyzing a subset of the answers to the COVID19Impact-Survey, a very large, anonymous, online citizen survey about COVID-19³⁴. Launched on March, 28th 2020 in Spain, the survey has since collected over 600,000 anonymous answers mainly from 4 countries (Spain, Italy, Brazil and Germany). From this very large sample, we analyze the data from Spain for the time period between April, 3rd 2020 and September, 11th 2020.

From an original set of 25 questions³⁴, we analyze the answers to the questions included in Table S1 in the Supplementary Information. Given that the data was collected by means of a non-probabilistic sampling method, we weigh the answers such that the distribution of answers per age, gender and geographic province matches the officially reported distributions in the Spanish census data. A similar methodology was implemented in³⁴. Table S2 in the Supplementary Information shows the distributions of the age and gender variables before and after weighting the survey answers. The Table also includes the official census data for comparison.

Table 1 depicts the dependent and independent variables used in our analysis. We are particularly interested in understanding the differences in demographic factors, impact and perceptions between those who report an opposition to be confined vs those who report a willingness to be confined for at least 6 months. Thus, we model our dependent variable (willingness to be confined) as a binary variable where 0 represents an unwillingness to be confined and 1 represents maximum acceptance towards confinement (willing to be confined for 6 months or more). From a total of 109,515 unique answers, we selected the answers of those reporting an unwillingness to be confined and those reporting an acceptance to be confined for at least 6 months, yielding a sample with 20,054 unique responses for the time period between April, 3rd and September, 11th 2020.

Table 2 shows the resulting proportions of the variables analysed in the study. As seen on the Table, the data set is unbalanced regarding the target variable: there are 7,887 (39.3%) answers of those reporting an unwillingness to be confined vs 12,167 (60.7%) answers of those who report high willingness towards confinement (≥ 6 months). In terms of gender, there are 8,237 (41%) and 11,817 (59%) answers by women and men, respectively. Note that the survey provides three options to the gender question: male, female, and another gender. However, in the time period of study, we only received 0.25% of the answers corresponding to "another gender", which did not allow us to make any statistically significant analysis for that option. Hence, it is not included in the analysis. We group the age of our participants into four possible age buckets: [18–29], [30–49], [50–59] and 60+, with 9%, 44%, 27% and 20% of the answers, respectively.

To consider the impact of time, we include a discrete variable called *Phase* that refers to the different phases of confinement that were applied in Spain in the spring and summer of 2020, as shown in Fig. 5 and described in Table 1. From Table 2, we observe a decrease over time in those reporting a willingness to be confined for a long time period: at the beginning of the pandemic (*Workplace closure* phase), they represent 85.8% of the sample, whereas once the *New normal* phase was declared at the end of June of 2020, this group was only 55.4% of the sample. Thus, we use the Phase variable to segment the sample in 8 different time periods, and we build models both for the entire time period and for each of the confinement phases independently.

Variable/Values	Explanation
Willingness to be confined	
0	Willingness to be confined for 0 days
1	Willingness to be confined for ≥ 6 months
Phase	
Workplace closure 28th March 2020	Closure of all non-essential labor
Workplace reopening 12th April 2020	Return to work in sectors affected by Workplace closure
Walks allowed 28th April 2020	Citizens are allowed to leave their homes to go for walks after domiciliary confinement
Phase 0 4th May 2020	Public transport/museums/churches at 33% capacity Takeaway food open
Phase I 11th May 2020	Phase 0 + Reopening of Hotels (except common areas) Restaurant terraces to 50% capacity Mobility inside the province to travel to second residences
Phase II 25th May 2020	Phase I + Reop. non-university classrooms Restoration/malls(non common areas)/churches at 50% Reop. beaches; Cultural spaces/ hotels common areas to 33%
Phase III 8th Jun 2020	Phase II + Reop. discos without dance floor at 50% capacity Public transport at 100% capacity Reop. hotels/malls common areas/cultural spaces at 50% capacity Mobility inside the autonomous region
New Normal (21st Jun 2020 onwards)	Normal activity is resumed; Masks are compulsory in public spaces
Age (Q1)	
"18–29"	18 to 29 years old
"30–49"	30 to 49 years old
"50–59"	50 to 59 years old
"60+"	60 or more years old
Gender (Q2)	
Man	Male respondents
Woman	Female respondents
Home (Q5)	
Apartment	The respondent lives in an apartment
Single family	The respondent lives in a single family house
Household (Q6)	
Young	Every house member is aged < 60 years old
Elderly	Without any children at home and every house member aged 60+
Multi-generational	With children (aged < 18) and members aged 60+
Economic Impact (Q15)	
None	No economic impact is reported
Mild	I lost my job, I lost part or all of my savings My business is in danger of bankruptcy My employer is at risk of bankruptcy
Severe	I can't afford to buy food, I can't pay my rent or mortgage anymore
Psychological Impact (Q25)	
None	No psychological impact is reported
Mild	Excessive use of technology by adults (tablet, phone, TV) Excessive use of technology by children (tablet, phone, TV)
Severe	High levels of anxiety, High levels of stress, High levels of loneliness High levels of sadness, Loud arguments or fights with other members of the home Excessive consumption of alcohol Excessive consumption of drugs (prescription or other)
CumInc	Average 14-day Cumulative Incidence per 100,000 people in each of the 17 Autonomous Regions in Spain

Table 1. Independent and dependent variables used in our study.

Variable	0 days	6 or more months
Willingness to be confined	7,887 (39.3%)	12,167 (60.7%)
Phase		
Workplace closure	55 (14.2%)	331 (85.8%)
Workplace reopening	626 (24.8%)	1,894 (75.2%)
Walks allowed	516 (36.9%)	883 (63.1%)*
Phase 0	426 (30.4%)	977 (69.6%)
Phase I	653 (40.0%)	979 (60.0%)*
Phase II	602 (43.1%)	794 (56.9%)
Phase III	377 (40.2%)	561 (59.8%)*
New Normal	4,632 (44.6%)	5,748 (55.4%)
Age		
18–29	885 (47.3%)	988 (52.7%)
30–49	3,549 (39.9%)	5,338 (60.1%)*
50–59	2,095 (39.1%)	3,263 (60.9%)*
60+	1,358 (34.5%)	2,578 (65.5%)
Gender		
Male	4,883 (41.3%)	6,934 (58.7%)
Female	3,004 (36.5%)	5,233 (63.5%)
Home		
Apartment	5,265 (41%)	7,579 (59%)
Single family	2,622 (36.4%)	4,588 (63.3%)
Household		
Young	5,529 (40.8%)	8,006 (59.2%)
Elderly	1,863 (34.6%)	1,629 (65.4%)
Multi-generational	1,495 (37.1%)	2,532 (62.9%)
Economic impact		
None	4,139 (68.6%)	9,036 (31.4%)
Mild	2,514 (47.5%)	2,273 (52.5%)
Severe	1,063 (34.9%)	571 (65.1%)
Psychological impact		
None	2,612 (25.0%)	7,822 (75.0%)
Mild	1,859 (47.2%)	2,081 (52.8%)
Severe	3,416 (60.1%)	2,264 (39.9%)

Table 2. Descriptive statistics of the subset of answers to the survey analyzed in this paper. All the variables pass a proportions test, with p-values < 0.001, except for the variables marked with *.

We first performed a proportions test to determine if the distribution of answers is significantly different for each of the variables of interest between those unwilling to be confined and those with high levels of acceptance towards confinement. We found statistically significant differences in all the variables (p-value < 0.01) except for the variables marked with * on Table 2, which correspond to *Walks allowed*, *Phase I* and *Phase III* and for the age groups 30–49 and 50–59. Nonetheless, note that all the variables have at least two values with significant differences. Thus, we conclude that all variables are relevant to be included in our models.

The main research questions that we would like to answer through the analysis of this data set are:

- (1) **RQ1:** Are there differentiating attributes between those who report high vs no willingness to be confined?;
- (2) **RQ2:** Is it possible to accurately predict an unwillingness to be confined from the independent variables? and
- (3) **RQ3:** How did the willingness to be confined evolve during the different phases of the confinement?

Logistic regression model. We obtain a logistic regression model with all the variables described in Table 1 and the interactions shown in Table 3. We refer the reader to the Materials and Methods section for a description of our approach. We compute the McFadden's pseudo-R squared (0.131), the Cox-Snell residual (0.161) and Nagelkerke's R squared (0.218) to assess the goodness of fit of the selected model.

As reflected in Table 3, the variables with the most interactions are *Gender* and *PsyI*. The existence of the variable *Phase* in our model, interacting with *PsyI*, *Gender* and *Household*, empirically corroborates a temporal dependency of the target variable. Interestingly, the 14-day Cumulative Incidence of COVID-19 cases (*CumInc*) has no significant interaction with any of the variables.

Variable	Significant interactions with
Phase	PsyI, Gender, Household
Gender	EcoI, Home, Phase, Household
Age	PsyI
Home	Gender
Household	PsyI, Phase, Gender
EcoI	PsyI, Age, Gender
PsyI	EcoI, Age, Household, Phase
CumInc	–

Table 3. Significant interactions identified by the logistic regression model.

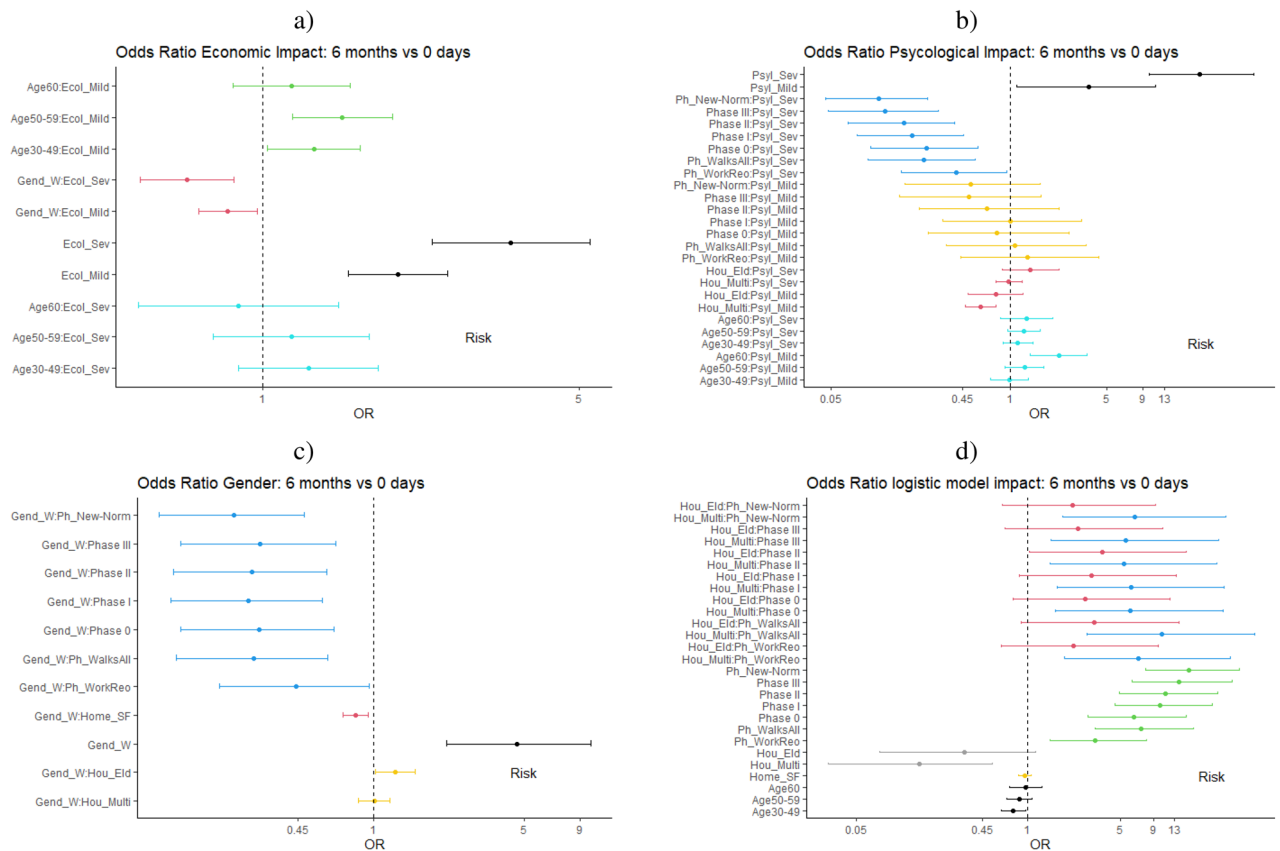


Figure 1. Odds ratios for all the variables with interactions in the logistic regression model.

We performed an Odds Ratio (OR) analysis to shed light on the role of each of the variables in the logistic regression model. A variable with an OR > 1 is typically interpreted as a risk factor, *i.e.*, it is a variable that significantly increases the probability of the target variable to be 1. Conversely, a variable with an OR < 1 decreases the probability for the target variable to be 1. Thus, we reverse the coding of the target variable so it represents a *risk*: we code as 1 the answers corresponding to an unwillingness to be confined and as 0 the answers corresponding to an acceptance towards a confinement of at least 6 months.

In our case, the basal categories for the OR analysis are: *Phase: Workplace closure; Gender: Men; Age: 18–29; Home: Apartment; Household: Young; EcoI: None and PsyI: None*, which correspond to the first row of each category in Table 2.

Note that the OR value in a logistic regression model with second order interactions among some of the variables (*Phase, Gender, Age, Home, Household, EcoI and PsyI* as per Table 3) is given by the OR of the main variable multiplied by the OR of the interaction variables where the main variable is present.

Figure 1 depicts the OR of the variables with the most important interactions, grouped in four categories to ease their interpretation.

Economic impact. Figure 1a shows the OR for the *EcoI* variable taking as basal category *None*. Note how the OR is larger than one for all values of the *EcoI* variable, meaning that this variable is a risk factor. Moreover, the OR

for *EcoI = Severe* (3.53) is larger than for *EcoI = Mild* (1.99) (in black color): those who report having had severe economic impact are more likely to report that they are not willing to be confined when compared to those with mild economic impact.

In addition, the OR of the interaction *Age:EcoI* (cyan lines for Severe and green for Mild) is larger than 1 for those aged 30–59 years old. Thus, people in that age group are more likely to report that they are not willing to be confined than those aged 18–29 years old (basal category). In the case of respondents aged 60+ years old, the OR is greater than one only when the economic impact is mild.

In terms of the *Gender:EcoI* interaction, the OR are less than 1 for both values of *EcoI* (red line). Hence, women with any kind of economic impact are less likely to report not willing to be in confinement any longer when compared to men (basal category).

Psychological impact. Figure 1b depicts the OR for the *PsyI* variable taking as basal category *None*. As shown in the Figure, those who report (*Severe or Mild*) psychological impact have higher risk to report an unwillingness to be confined than those without any psychological impact. Moreover, the larger the psychological impact, the higher the risk (black lines). In terms of the interaction *Age:PsyI*, all the age groups show a similar level of risk (cyan lines), except for those aged 60+ years old with mild psychological impact who seem to be at significantly higher risk of reporting that they would not be able to be in confinement any longer (OR = 2.23). We observe a difference in the risk depending on the age of the members of the *Household*: households composed of elderly (60+) are significantly more at risk to report an unwillingness to be in confinement when compared to the rest (red and cyan lines). Concerning the interaction between *Phase:PsyI*, it seems that time has an attenuating effect on the risk as there is a consistent decrease of the risk over time. These results occur for *PsyI = Severe* (blue lines) and *PsyI = Mild* (yellow lines) but in this case to a lesser extent. This finding is surprising, as one would expect people to be at higher risk of not tolerating confinements the longer the pandemic lasts.

Gender. In Figure 1c the OR for the *Gender* variable are shown, taking *Men* as the basal category. As depicted in the Figure, women tend to be at significantly higher risk (OR = 4.60) than men to report that they are not able/willing to be in confinement any longer (black line). The value of this OR is significantly larger than the OR found for the *EcoI* and *PsyI* variables. This risk is increased for women who live in households composed of elders *Household_{Elder}* (yellow lines). In terms of the interaction *Phase:Gender*, all the OR are similar and lower than 1 (blue lines), meaning that as time goes by, the risk for women to report that they are not able to be in confinement is lower than during the reference phase (*Workplace closure*).

Rest of variables. Figure 1d depicts the OR of the rest of variables and their interactions. As seen in the Figure, the OR of *Phase* as a main factor is larger than one for all its values when compared to its basal value (*Workplace closure*) and increases over time from 3.3 during the *Work reopening* phase to 16.7 during the *New normality* phase. In other words, the risk of reporting an unwillingness to stay in confinement increases significantly as time goes by.

Regarding the type of *Household*, multi-generational households are at lower risk of reporting that they are not willing to be confined any longer than households with composed of elders (grey lines). However, in both cases the OR < 1, so they are attenuating factors when compared to the basal category (*Household = Young*).

We do not observe a significant impact of *Age* as a main factor on the risk to report being unwilling to be confined any longer. Nonetheless, as previously reported, age can be a risk factor in its interaction with other factors, such as *Economic* and *Psychological Impact*.

Combining all the OR reported above, we conclude that men living in multi-generational homes, with neither psychological nor economic impact due to the pandemic, are the least likely group to report an unwillingness to be confined. Conversely, women aged 30+ years old with severe economic and psychological impact and living in households with young children are at the highest risk to report an unwillingness to be confined.

Classification performance. Given that we have a quantitative model of the unwillingness to be confined, we test the model as a binary classifier to automatically infer such unwillingness from our independent variables. We randomly split the data in a training (80% of the data) and testing (20% of the data) sets. Note that we group the data on a weekly basis to have enough samples of survey answers across the entire time period of study.

On this binary classification task, the logistic regression model has a sensibility of $72.5\% \pm 0.73$ and a specificity of $63.6\% \pm 0.75$ with 100 random permutation data for training. A random model would have a sensibility of $50.0\% \pm 0.004$ and a specificity of $49.99\% \pm 0.005$. We experimented with other state-of-the-art machine learning methods. Gradient Boosted Trees (GBTs) trained following the same sampling methodology obtained a sensibility of $58.0\% \pm 0.01$ and a specificity of $60.7\% \pm 0.01$. The three most important variables identified by the GBTs were *Psychological impact*, *Phase* and *Economic impact*.

Automatic pattern extraction. In addition to the logistic regression modeling, we have developed an automatic pattern extraction algorithm described in the Materials and Methods Section. Using this model, we identify the most influential variables and discover the most recurrent patterns among those who report an unwillingness to be confined.

Most influential variables. Figure 2 depicts the variable importance obtained with our pattern extraction methodology across the different phases of the pandemic. As displayed in the Figure, there is a significant change over

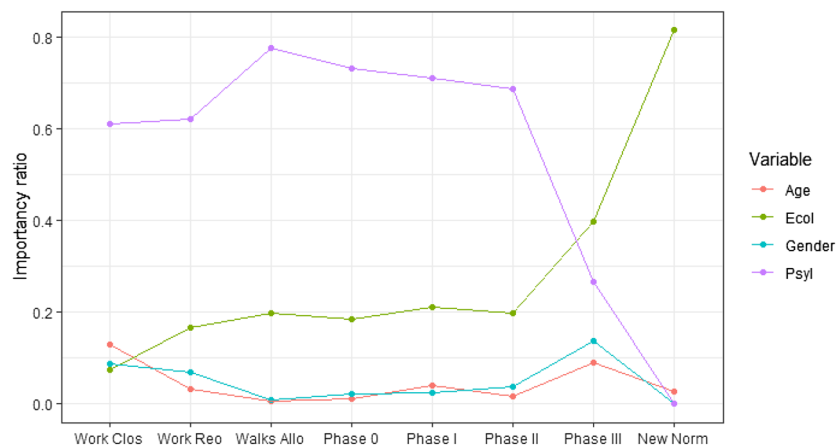


Figure 2. Evolution of the variable importance over time, according to the pattern extraction model.

time of the importance of two variables: the *Psychological (Psyl)* and *Economic Impact (EcoI)* of the pandemic on the survey respondents.

Initially, *Psyl* is the most important variable, increasing its importance up to the *Walks allowed* phase, when its importance begins to decline. Conversely, the *EcoI* variable is not important at the beginning of the pandemic. Up until the end of *Phase II*, the importance of this variable increases smoothly or remains constant over time. However, with the start of *Phase III* and the return of the new normality, the importance of the economic impact to determine the unwillingness to be confined becomes evident.

The curves corresponding to the importance of the psychological and economic impact variables intersect between *Phase II* and *Phase III*. The rest of the variables exhibit a fairly stable behaviour over time, with mild fluctuations in their importance, affecting the target variable almost uniformly throughout the different phases of confinement.

Given this dependency of the target variable on the Phase of the pandemic, we discuss next our findings when performing an analysis in each of the 8 Phases of study (see Fig. 5).

Analyses per phase of the pandemic. *Logistic regression.* Figure 3 depicts the OR of the main effects and their temporal evolution over time in each of the phases of the pandemic. Note that for each *Phase* we build a different logistic model with the main variables but without interactions given that we do not have enough data per Phase to identify meaningful interactions.

Figure 3a shows the Odds Ratio of the 14-day cumulative incidence over time. The OR increases over time, which might be indicative of the so called *pandemic fatigue*. In the early phases of the pandemic, the OR is less than one, meaning that the larger the 14-day cumulative incidence, the lower the probability for respondents to report an unwillingness to be in confinement. Conversely, from *Phase II* onward, respondents seem to be indifferent (OR close to 1) to changes in the cumulative incidence. Figure 3b depicts the OR of the *Psychological Impact (Psyl)* variable taking as basal category *No impact*. As seen in the Figure, the OR are always larger than 1 and decrease over time. Hence, the role that the psychological impact plays in increasing the probability of reporting not being able to be confined any longer decreases as the months go by. This finding is consistent with the previously reported variable importance analysis (see Fig. 2).

In the case of the *Economic Impact (EcoI)* variable (Fig. 3c), the OR are also larger than 1 in all the phases. However, the evolution over time is different than that of the *Psyl* variable. We observe a significant increase in the OR after *Phase III*, particularly for those reporting severe economic impact. This result is also consistent with the variable importance analysis previously described. Moreover, it makes intuitive sense: as the economic situation of people worsens due to the pandemic, their probability of reporting that they are not willing to be in confinement any more increases.

The changes over time in the OR for the *Age* variable are depicted in Fig. 3d. Those aged 60+ are at significantly larger risk in the first phases of the pandemic to say that they “can’t anymore” when compared to the rest of age groups. Once the *New normal* phase started at the end of June of 2020, the risk of those aged 50–59 and those aged 60+ becomes similar. Interestingly, the 30–49 age group is the only age group with an OR < 1 during the entire period.

Those living in single family homes are at lower risk to report that they cannot be in confinement any longer when compared to those living in an apartment (Fig. 3e). This finding makes intuitive sense as single family homes are typically more spacious than apartments in Spain. Note that according to official data from the Spanish National Institute of Statistics, 57% of Spanish households consist of apartments of 90 square meters or less.

Finally, *women* are at larger risk than men to report that they cannot endure the confinement any longer throughout the entire period until reaching the *New normal* phase, when the OR for women becomes smaller than 1 Fig. 3e.

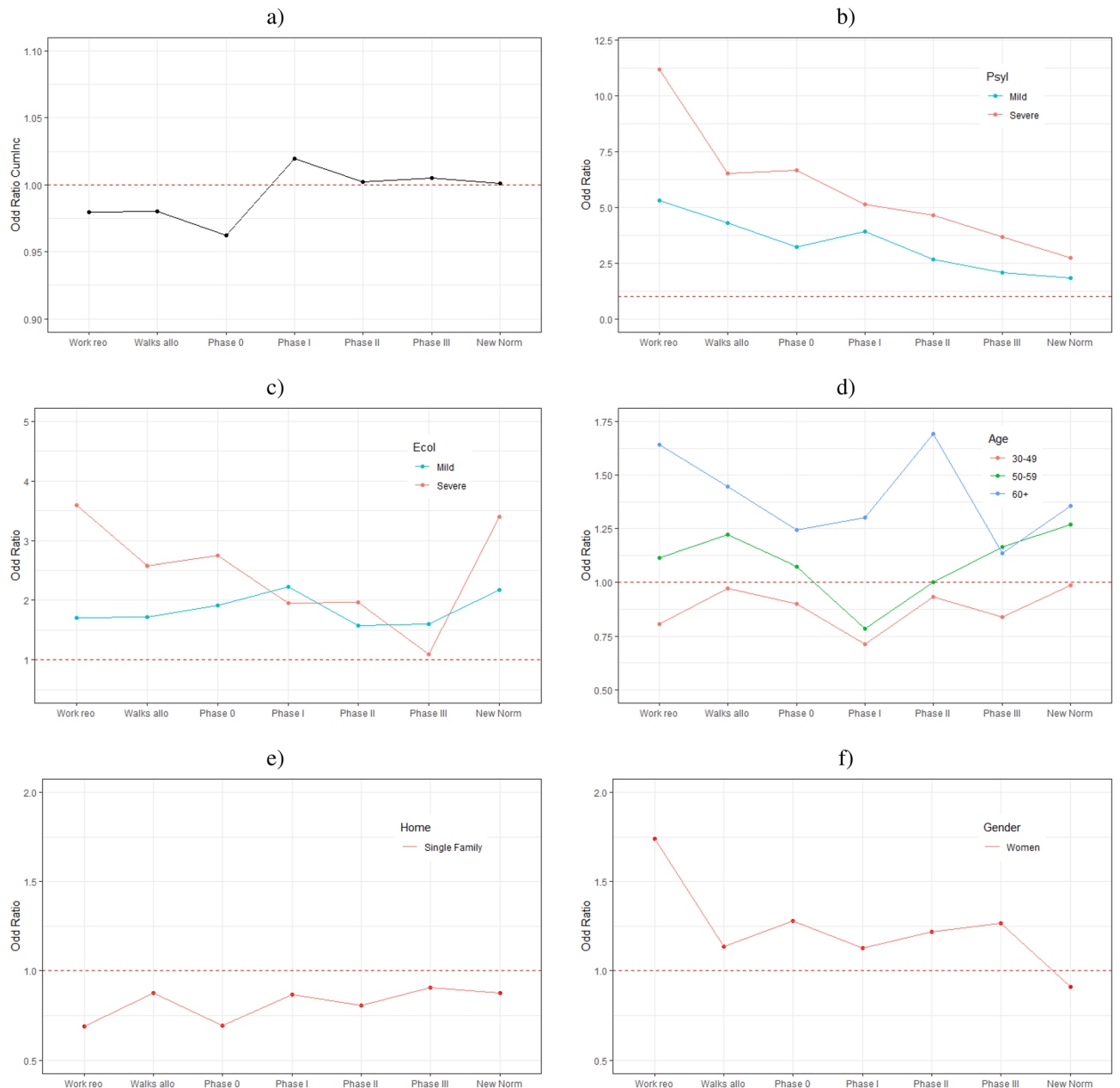


Figure 3. Odds ratios for the main variables of the models per phase. (a) CumInc (b) Psychological Impact, (c) Economic Impact, (d) Age, (e) Type of Home and (f) Gender. Each point on the X axis corresponds to the OR of the logistic model for that particular phase.

Automatic pattern detection. Given the changes of the most relevant variables over time (see Figs. 2, 4), we apply our automatic pattern extraction algorithm to the data in each of the phases. We extract the patterns that model the unwillingness to be confined from the two most influential variables: economic and psychological impact, together with gender and age.

Figure 2 shows a graph comparison of the weights of the most important variables as per our pattern detection method applied in each of the phases of confinement. As seen in the Figure, each confinement phase is modeled differently by these explanatory variables. Thus, the patterns identified in each phase change as the variables become more or less important.

As shown in Fig. 4d, age and gender have relatively important weights in the first stage of confinement but then practically disappear until Phase III to disappear again in the New normal phase.

The intermediate phases of confinement (from Walks allowed to Phase II) can be mainly modeled using the Psychological Impact variable. In Phase III (the last stage prior to the New normal), the unwillingness to be confined is affected by the two socio-demographic variables (gender and age), by psychological impact and especially by economic impact, which is triggered in this phase. In the New normal phase, the unwillingness to be confined is mainly dependent on economic impact.

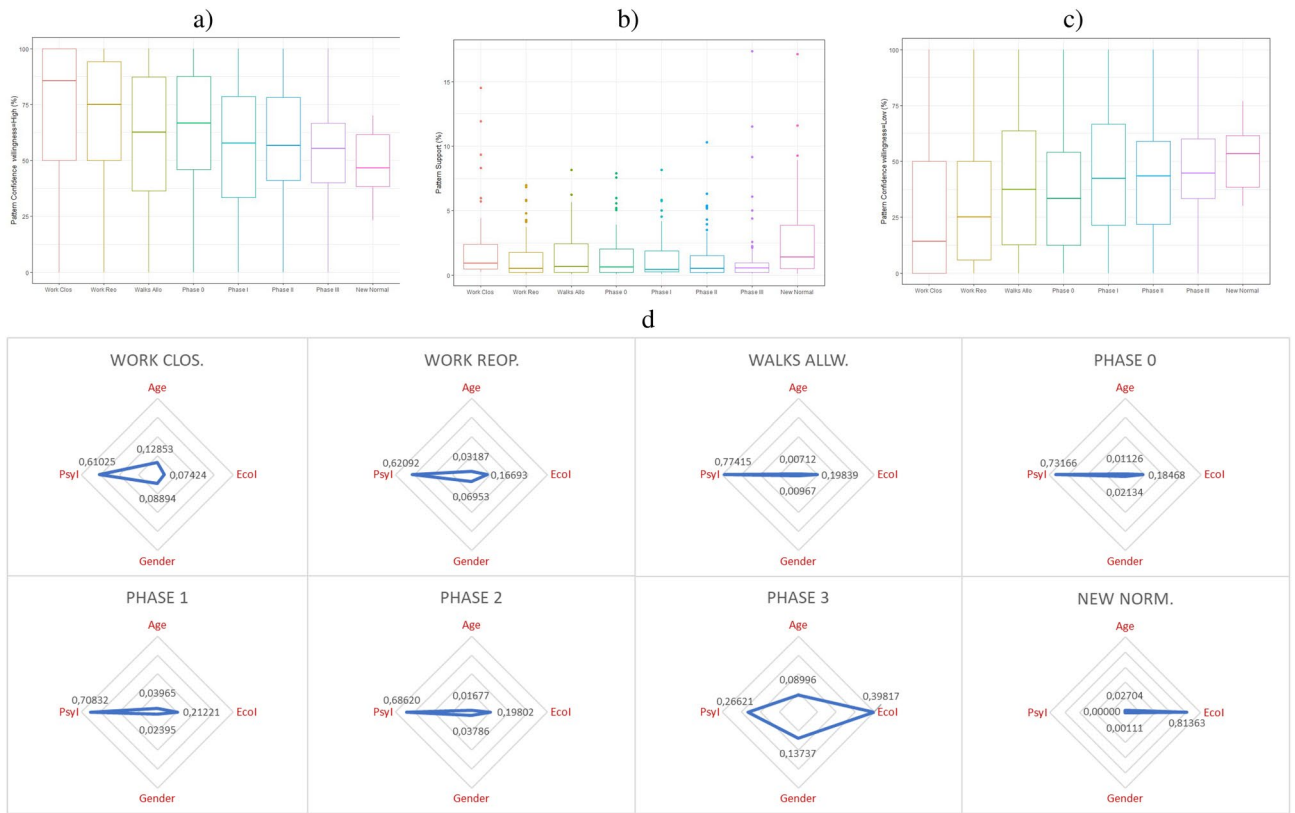


Figure 4. (a) Confidence statistics of the patterns in each phase for the willingness to be confined for 6 months or longer. (b) Support statistics of the identified patterns in each phase. (c) Confidence statistics of the patterns in each phase for the willingness to be confined for 0 days. (d) Weights of the most important variables according to our pattern detection method per confinement Phase.

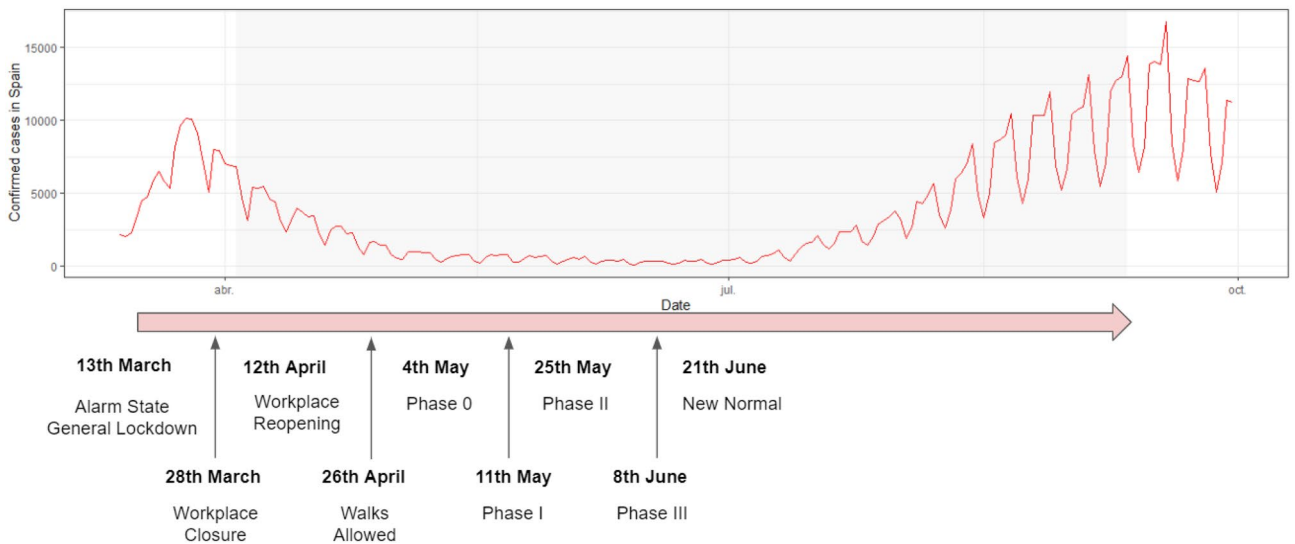


Figure 5. Confinement phases and daily number of confirmed COVID-19 cases in Spain between March and September of 2020. The shaded part of the graph corresponds to the time period of analysis.

Next, we summarize the most relevant patterns identified in each of the phases.

1. Workplace closure:

In this phase, those over 50 years of age, regardless of their gender and economic impact, as long as they report no psychological impact (26.4% of the sample) always opt (100% of the time) for a willingness to be confined for 6 months or longer.

While we did not obtain any significant interaction between gender and psychological impact in our logistic regression modeling, we did identify significant gender differences in the 30–49 years old age range: among those who do not report any economic impact but do report severe psychological impact (close to 10% of the sample): 82.6% of men vs 64.7% of women would be willing to be in confinement for 6 months or longer. This finding illustrates the complementary nature of our modeling approaches. The pattern discovery method identifies patterns that involve three or more variables, which would be very difficult to achieve via our logistic regression methodology.

In this initial stage of confinement, when the psychological and/or economic impact might not yet be evident, the willingness to be confined is at its highest levels. However, 2.6% of the sample corresponds to women between 50 and 59 years old who report severe psychological impact. Of these, 60.0% report an unwillingness to be in confinement, even in this very early stage of the pandemic.

2. Workplace reopening:

A large percentage (close to 94% on average) of men and women between 30 and 60 years old, who report neither economic nor psychological impact (18.4% of the sample) are willing to be confined for 6 months or longer. When this group reports mild psychological impact (13.1% of the sample), we observe gender differences, with 92.5% of men vs 85.4% of women willing to remain in confinement for 6 months or longer. As expected, the reopening of workplaces hardly had any impact on respondents aged 60+ (mostly of non-working age) without economic or psychological impacts: 94.1% of them on average report a willingness to be confined for 6 or more months, irrespective of their gender. This group represents 11.6% of the sample.

At this stage, those who report high levels of acceptance towards confinement still clearly outweigh those unwilling to be confined. However, once again, we identify clear gender differences in the patterns, with women being more likely to report that they cannot be in confinement any longer when compared to men. In particular, we identify two distinct patterns: first, 43.6% of women between 50 and 59 years old who report severe psychological impact and without economic impact (3.7% of the sample) report minimal acceptance towards confinement; second, 56.5% of women between 30 and 49 years old, who in addition to reporting severe psychological impact also claim to have medium economic impact (2.7% of the sample) also respond that they would not be able to continue in confinement.

3. Walks allowed:

During this phase, the participants' opinions regarding their willingness to be confined starts to balance: the percentage of those willing to be confined for 6 months or longer decreases whereas those who are not willing to be confined increase.

The first identified pattern corresponds to people between 30 and 49 years old who report having no economic impact but severe psychological impact (13.0% of the sample). Among these (men and women alike), 54.8% on average report a willingness to be confined for 6 months or longer.

At this stage, there are several groups (both men and women) that begin to opt mainly for an unwillingness to be confined. The pattern with the largest support reveals that 75.0% of women between 30 and 49 years old with medium economic impact and a severe psychological impact (3.7% of the sample) report an unwillingness to be confined.

4. Phase 0:

In the first phase of the re-opening in Spain, more people report being able to continue in confinement for 6 months or longer than those reporting that they cannot stand it any longer, with the exception of two groups: women between 30 and 49 with severe psychological impact and mild economic impact (3.49% of the sample) and women over 60, with severe psychological impact but no economic impact (6.0% of sample). In these two groups, only 46.9% and 45.2% respectively report a willingness to be confined for 6 months or longer.

While the willingness to be in confinement for 6 months or longer also decreases among men, still remains the most popular option for them. This finding is aligned with our OR analysis shown in Fig. 1c.

Half of those aged 60+ (both men and women) who do not declare to have any economic impact but report severe psychological impact (9.0% of the sample) report not being able to remain in confinement any longer. The burden of the pandemic starts to become evident.

5. Phases I and II:

Both phases show similar patterns where the unwillingness to be confined is prevalent between individuals who report severe psychological impact with some economic impact. These levels of unwillingness are higher when, in addition to the severe psychological impact, participants also report some type of economic impact. In Phase I, the opinion of some groups is polarized clearly towards an unwillingness to be confined: 67.2% of women between 30 and 59 years old who report mild economic impact and severe psychological impact (over 7.4% of the sample) and 86.9% of women aged 60+ who, having no economic impact declare severe psychological impact (5.0% of the sample) respond that they cannot stand the confinement any longer. Again, the most significant patterns related to those who are unwilling to continue in confinement are found amongst women.

6. Phase III:

In the last phase of the re-opening and before the imminent start of the *New normal* phase, we observe an increase among those who report being able to remain in confinement for 6 months or longer. As seen in our feature importance analysis, the psychological impact becomes less relevant than economic impact and demographic factors. Hence, psychological impact does now appear in most of the patterns.

In general terms, the majority opinion is once again towards a willingness to be confined for 6 months or longer. The group that reports the highest levels of unwillingness to be in confinement (47.3% of the time) are people (men and women) between 30 and 49 years old who do not report psychological impact but report economic impact (7.6% of the sample).

7. New Normal:

The willingness to be confined shows a balanced distribution between those who affirm that they are not able to stand it any longer and those who would be willing to be confined for 6 months or longer: 65.2% of people between 30 and 50 years old (without distinction of gender) who report neither economic nor psychological impact (28.8% of the sample) would accept long-term confinement, compared to 34.8% who would not stand confinement any longer.

During this period (between the end of June 2020 and September 2020), the economic impact is the most influential factor to determine the willingness to be confined, when compared to gender, age and psychological impact. We highlight two patterns that strongly support (in more than 67.7% of the cases) an unwillingness to be confined: women aged between 30 and 49 years old who do not report psychological impact, but report severe economic impact (3.4% of the sample), and men aged between 50 and 59 years old, who do not report psychological impact, but who report moderate economic impact (2.7% of the sample). Both groups are of working age and with some level of economic impact. They clearly opt for an unwillingness to be confined.

Discussion

We have described two different analytical approaches applied to the same data. The obtained results are consistent yet also complementary with each other. The logistic regression modeling revealed the main factors that impact the unwillingness to be confined and enables us to build a predictive model of such an unwillingness. The pattern discovery methodology identified relevant patterns that characterize the groups of respondents who exhibit an unwillingness to be in confinement.

From these analyses, we draw several key findings related to the interplay between demographic factors, the economic and psychological impact of the pandemic on people's lives and the passage of time.

1. Pandemic fatigue is evident Our sample spans over a 5-month period which includes both the severe lockdown imposed by the Spanish government between March 2020 and June 2020 and more than two months of the *New normal* phase. As the weeks go by, we observe a clear decrease of those who report a willingness to be confined for 6 months or longer from 85.8% to 55.4% of the sample. We corroborate this finding in our OR analysis, where the risk to report an unwillingness to be in confinement any longer significantly increases with the *Phases* of the pandemic (see Fig. 1d). The most important variables to determine the unwillingness to be confined change as the months go by: psychological impact starts as the most important variable whereas economic impact becomes the most important variable once the *New normal* phase starts.

Interestingly, until *Phase 0* of the pandemic, the larger the 14-day cumulative incidence of COVID-19 cases, the lower the risk of reporting an unwillingness to be confined. However, as the months went by, the impact of the cumulative incidence on the unwillingness to be in confinement disappeared, starting in *Phase II* of the pandemic.

Since the World Health Organisation published its report on pandemic fatigue³⁵, several authors have investigated this topic. Reicher and Drury propose that the non-adherence to the confinement measures is not due to pandemic fatigue, but rather due to structural weaknesses in our society, such as the inability to effectively self-isolate³⁶. Conversely, other authors have written about the psychological, physical and economic impact of the pandemic on people, which has led to the emergence of the pandemic fatigue in different countries in the world^{37–42}.

Given that the SARS-CoV-2 pandemic is expected to last for many more months, new confinements might be needed to control the transmission rate of the disease. However, the population is in a very different mental and economic state than they were at the beginning of the pandemic. Hence, policy makers should be sensitive to this situation in order to maximize the chances that future confinement measures will be both accepted and complied with by the population.

2. Gender and age matter We identify very significant gender differences in the reported unwillingness to be confined, with women being more likely to be unwilling throughout all the phases of the pandemic. According to the OR analysis, age does not seem to play a significant role in increasing the risk to report an unwillingness to be confined. However, our pattern discovery method identifies several patterns involving individuals (and especially women) of working age with some level psychological and/or economic impact and a clear unwillingness to remain in confinement any longer.

Previous work has also identified significant gender differences, with women being generally more compliant with individual protection and confinement measures^{43,44} than men. Galasso et al. report that women are more likely to perceive the pandemic as a very serious health problem and to agree and to comply with restraining measures from the analysis of survey data in eight countries conducted in March–April 2020⁴⁵. Haischer et al.⁴⁶ studied facial mask wearing behavior in a sample of 9,935 shoppers and identified that women were 1.5 times more likely to wear a mask than men. The fact that women in our study are at larger risk of not being willing to be in confinement when compared to men might be explained by the fact that women also are more likely to report psychological impact. According to our data, 18.0% of men vs 20.8% of women and 25.0% of men vs 30.6% of women ($\chi^2 = 143.45$, p -value $< 2.2e-16$) report mild and severe psychological impact due to the pandemic, respectively. Previous research has also studied the differential psychological impact of the pandemic on women^{47–49}.

In addition to the increased risk to report psychological distress, there are other factors that might have contributed to the identified gender-based differences in the reported willingness to be in confinement. First, we have evidence from the survey (Q21 described in³⁴ and visualization available in⁵⁰) that women in Spain aged 30–59 are more than twice as likely as men in the same age range to report not being able to self-isolate due to childcare duties (26% of women vs 11% of men). Other studies have also found that the burden of

the household and childcare work has been primarily carried by women during the pandemic⁵¹. Second, we cannot obviate the impact of gender violence. Recent work has reported an increase of gender violence in the homes in Spain during the pandemic⁴⁹. An indicator of this increase in gender violence is the number of calls made to the gender violence hot-line in Spain (number 016 provided by the Ministry of Equality), which significantly increased from a average of approximately 5,500 calls per week before the lockdown to over 8,500 calls during the lockdown period. Finally, recent work has found that the COVID-19 pandemic increased gender inequality both in paid and unpaid work in Spain⁵².

We believe that these factors—namely, the burden of childcare and house-related work, gender violence in the home and differential labor impact—are key contributors to increasing the psychological distress and the unwillingness by women to stay in further confinement when compared to men.

These gender and age-based differences are important to policy makers, as they reveal the need for gender and age-dependent public policies that would address the needs of different gender and age groups.

An example of such policies are the Spanish government's policies related to gender violence. Anticipating an increase in the number and intensity of gender violence incidents, the Spanish Government implemented a series of measures to protect women during the lockdown. The government approved the Royal Decree 12/2020 and considered all services that protect and assist victims of gender-based violence as essential, such that they were operative and even were intensified⁵³ during the entire lockdown period. The services included a telephone hot-line of attention and protection for victims of gender violence and an electronic monitoring system to enforce restraining orders.

Despite such measures, the negative unequal impact of the pandemic on women's lives in Spain—including increased levels of gender violence—has been reported and documented. Beyond the necessary cultural change and the investment in education programs, additional policies could include providing free housing for those at risk of gender violence in the home, social and psychological support, extended and accessible childcare services (even during lockdowns), and self-isolation infrastructures to ease the burden of those in need of quarantine.

3. From psychological to economic impact Psychological impact played a very important role in the early stages of the pandemic (up to *Phase III*): it is the main variable that determines the unwillingness to be confined, as reflected in the reported OR and feature importance analyses. However, as time progressed, psychological impact became less relevant whereas the economic impact of the pandemic became the main driver to determine an unwillingness to be confined. In the *New normal* phase, those with severe economic impact are significantly more likely to report that they cannot stand the confinement any longer when compared to those without any reported economic impact.

Note that when the lockdown was first announced in Spain, it was expected to last only for 2 weeks, which meant that the economic impact of the measures would be temporary and short-term. Hence, it makes intuitive sense that the role of the economic impact on people's willingness to be confined was limited in this initial phase. As the weeks (months) progressed and people realized that the pandemic was a long-term challenge, the unwillingness to be confined due to economic reasons significantly increased given that many people ran out of savings and did not have an economically sustainable way to continue in lockdown. According to the survey, the percentage of those who reported running out of their savings due to the pandemic changed from 7% in March 2020 to 16% in August 2020. Economically, the most affected sector in Spain was the hospitality sector, which represents 12.4% of the Spanish GDP. According to the Spanish National Office of Statistics, the hospitality sector's revenues in 2020 went down by 50.1% reducing its workforce by 13.3%. In addition, 82.74% of the Spanish companies are micro-companies with 2 or less workers who received limited financial aid from the Government other than unemployment benefits.

Other authors have studied the psychological impact of the COVID-19 pandemic on different populations, including students⁵⁴, healthcare workers⁵⁵, and the general population^{56–58}. The economic impact of pandemics throughout history and for the COVID-19 pandemic has been analysed by several authors, including Ceylan et al.⁵⁹ who conclude that unemployment with its income effects will be the most prominent economic consequence of the pandemic. However, we have not found any previous work that has studied the evolution over time of the role played by both the psychological and economic impact on the willingness to be in confinement.

This finding points out the importance of deploying not only financial support but also psychological support measures as people's livelihoods are severely impacted by over one year of the coronavirus pandemic.

Conclusions

Population confinements have been one of the most widely adopted non-pharmaceutical interventions (NPIs) implemented by governments across the globe to help contain the spread of the SARS-CoV-2 virus. While confinement measures have been proven to be effective to reduce the number of infections, they entail significant economic and social costs. Thus, understanding the factors that determine the willingness of individuals to be confined during a pandemic is of paramount importance, particularly, to policy and decision-makers.

In this paper, we have studied such factors by means of a large-scale, online population survey deployed in Spain, one of the most affected countries in Europe by the pandemic. We consider socio-demographic, economic and psychological factors, together with reported 14-day cumulative incidence per 100,000 inhabitants. We analyze 109,515 answers provided during a 5-month period (between April 3rd, 2020 and September 11th, 2020) to shed light on the impact of the passage of time.

We apply both logistic regression and automatic pattern identification methodologies to determine the role that the factors play in determining the unwillingness to be confined and their evolution over time. We find evidence of *pandemic fatigue*: the unwillingness of people to be confined increases as the months go by. We

identify significant gender and age differences, placing women—and particularly working-age women—at a higher risk than men to be unwilling to be confined. Finally, we uncover the importance of psychological impact at the early stages of the pandemic to be replaced by economic impact once the *New normal* was declared at the end of June of 2020.

Our results point to the need to design age and gender specific public policies, to implement economic programs to support the many citizens whose livelihoods have been severely impacted by the pandemic and to address the evident pandemic fatigue as the success of potential future confinements will depend on the population's willingness to comply with them.

Collecting a large sample of survey data does not come without limitations. First, respondents have to be adults (at least 18 years old). Second, there might be a difference between the reported willingness to be in confinement and the actual compliance with the confinement measures. Third, there is a self-selection and sampling bias as the survey is filled out by volunteers who have learned about the survey via social media, WhatsApp, newspapers articles or Facebook ads, and who need access to a computing device (smartphone, tablet, PC) with internet connection. We mitigate this bias by weighing the raw data and by deploying gender-balanced Facebook advertisement campaigns.

Material and methods

The COVID19 impact survey. We report the results of analyzing a subset of the answers to the COVID19ImpactSurvey, a very large, anonymous, online citizen survey about COVID-19³⁴. Launched on March 28th in Spain, the survey has since collected over 600,000 anonymous answers mainly from 4 countries (Spain, Italy, Brazil and Germany). From this very large sample, we analyze the data from Spain for the time period between April, 3rd 2020 and September, 11th 2020.

The COVID19ImpactSurvey was launched in collaboration with the Valencian Government in Spain and has been regularly used to assist in their decision-making³⁴. The survey has changed over time to adapt to each of the phases of the pandemic. During the period of our study, the survey consisted of 25 questions designed to collect information about the participants' demographic and household situation (Q1–Q7); their social contact behaviour (Q8–Q12); their support for the government measures deployed to contain the spread of coronavirus (Q13) and their willingness to continue in confinement (Q14); the economic and labor impact of the pandemic in their lives (Q15–Q19); their health risk factors (Q20), their ability to self-isolate if needed (Q21), the presence of coronavirus-related symptoms (Q22), the duration of the symptoms (if any) (Q23), whether they were tested for coronavirus (Q24); and the psychological/emotional impact of the pandemic in their home (Q25).

All methods described in this manuscript were carried out in accordance with relevant guidelines and regulations. The overall study, the research protocol, and the survey instrument were approved by the Ethical Committee of the University Miguel Hernández, which is called the Office for Responsible Research (OIR). OIR is the official body overseeing and regulating the experimental procedures and research activities to offer guarantees that they are conducted following the appropriate ethical and legal framework. The approval code given to the project was AUT.DEM.XBV.06.20. Informed consent was obtained from all subjects who had to confirm being at least 18 years old to be able to participate in the survey.

We refer the interested reader to³⁴ for a detailed description of the survey questions, the weighing methodology applied to reduce biases in the data and the results of analyzing the first week of answers. Note that question Q25 related to the psychological impact of the pandemic was added on April 3rd and thus it is not included in the analysis presented in³⁴.

From the questions described above, we report the results of analyzing the subset of questions depicted in Table S1 in the Supplementary Information, namely Q1–Q6, Q14, Q15 and Q25.

Given that we have data for several months, we analyze the survey answers according to the different phases of confinement implemented in Spain between April and September of 2020. Figure 5 illustrates the chronology of each of the phases as per official data (See, for example⁶⁰), together with the daily number of confirmed COVID-19 cases. The shaded portion of the graph corresponds to the time period of analysis in this paper.

Data processing. Table 1 summarizes the independent and dependent variables used in our study. Our target or dependent variable is the willingness to be confined, captured by Q14. Our independent variables are the socio-demographic, economic and psychological impact measures, captured by questions Q1–Q6, Q15 and Q25, the confinement Phase as shown in Fig. 5, and the 14-day Cumulative Incidence of SARS-CoV-2 per 100,000 inhabitants averaged over the 17 Autonomous Regions in Spain.

To minimize biases in the data, we weigh the raw answers by gender, age and province, such that the resulting distribution of answers matches the distributions reported by the official Spanish census. We apply the same methodology as described in³⁴. Table S2 in the Supplementary Information depicts the age and gender distributions of the raw and weighted data, together with the officially reported census distributions.

In addition, we simplify several of the answers as follows:

1. Q4. We only consider two types of homes: apartments and single family homes. These two answers represent 94% of the raw answers.
2. Q15. We group the answers to the economic impact question into three levels as follows: (a) *None (0)*: No economic impact is reported; (b) *Mild (1)*: “I lost my job” or “I lost part or all of my savings” or “My business is in danger of bankruptcy” or “My employer is at risk of bankruptcy”; (c) *Severe (2)*: “I can't afford to buy food” and/or “I can't pay my rent or mortgage anymore”.
3. Q25. We group the answers to the psychological impact question into three levels: (a) *None (0)*: No psychological impact is reported; (b) *Mild (1)*: “Excessive use of technology by adults (tablet, phone, TV)” and/or

Antecedent (explanatory variables)						Consequent (target variable values)				
SUPP	SUPP%	Gender	Age	EcolImpact	PsycholImpact	>>	High	CONF%	Low	CONF%
1779	17,14	female	30-49	None	None	>>	1152	64,76	627	35,24
1205	11,61	male	30-49	None	None	>>	792	65,73	413	34,27

SUPPORT: Probability that antecedent is given

CONFIDENCE values for extreme willingness to be confined

Figure 6. Example of two of the patterns identified by our pattern extraction method.

“Excessive use of technology by children (tablet, phone, TV)”; (c) *Severe* (2): “High levels of anxiety” and/or “High levels of stress” and/or “High levels of loneliness” and/or “High levels of sadness” and/or “Loud arguments or fights with other members of the home” and/or “Excessive consumption of alcohol” and/or “Excessive consumption of drugs (prescription or other)”.

- Q14. We divide the target variable, *i.e.*, the willingness to be in confinement into two categories, as per the answers to question Q14: (a) *Unwilling* (coded as 0): Willingness to be confined for 0 days, the person cannot stand it anymore; (b) *Long-term* (coded as 1): Willingness to be confined for at least 6 months.

Logistic regression models. We built an explainable logistic regression model⁶¹ to predict the unwillingness to be confined (binary variable) using the variables described in Table 2—namely, *Phase*, *Age*, *Gender*, *Home*, *Household*, *Economic Impact (Ecol)*, *Psychological Impact (PsyI)* and *14-day COVID-19 Cumulative Incidence (CumInc)*—as our independent variables with plausible interactions among them.

Note that in this regression model, we reversed the coding of the target variable, *i.e.* we coded with a value of 1 the willingness to be confined for zero days (*i.e.* the unwillingness to be confined), and with a value of 0 the willingness to be confined for at least 6 months. We used this reverse coding, such that our target variable represents the *risk* for respondents not to accept a confinement, thus being consistent with the risk analyses that are typically performed in logistic regression modeling via an Odds Ratio (OR) methodology.

First, we built a complete model that included all the variables and interactions between them. Next, we performed feature selection by means of a bidirectional step-wise variable selection process, such that at each step we added or removed a main variable or an interaction according to the Akaike Information Criterion (AIC)⁶². Finally, we computed the p-value of all the remaining variables. If the p-values were not lower than 0.05 we performed a likelihood ratio test⁶³ between the model with the variable included and the model without the variable to determine whether to include such a variable in the final model or not.

We also used this model to automatically predict the target variable. Moreover, we carried out an Odds Ratio analysis to shed light on the role of each of the variables to determine the opposition to confinement, as described in the Results Section.

Automatic pattern extraction. In addition to the logistic regression analysis, we applied a pattern extraction method to automatically identify meaningful patterns that would differentiate the behaviours, impact and characteristics of those not willing to be confined vs those willing to be confined for a long time period.

We followed a two-step methodology to automatically extract the patterns from the survey answers:

- Step 1: Selection of the most influential variables.** We identified the most influential variables with respect to our target variable (*i.e.*, the willingness to be confined). Note that not all the variables in our study might be relevant when modeling people’s willingness to be confined. In this variable selection step, we used the Entropy-based Filters algorithm from the FSelectorRcpp library in R, which is based on the information gain, gain ratio and symmetrical uncertainty metrics⁶⁴.
- Step 2: Automatic pattern generation,** via a modified version of the RBS algorithm⁶⁵, a classifier based on the ID3 family⁶⁶. The RBS algorithm is an iterative method that, without building the full structure of a tree, automatically identifies a set of rules (patterns) from a data set of discrete variables. The subsequent ordering and filtering of the patterns depends on the *significance* of the rule, which is a metric based on the classic concepts of support and confidence of the classification rules, but defined by intervals. We adapted the RBS algorithm to be used with re-weighted input samples rather than absolute supports, as it is usually done.

As a result of this two-step process, we obtained the set of rules or patterns that best model our target binary variable (willingness or acceptance to be confined). Each of the automatically identified patterns or classification rules has the structure shown in Fig. 6. The Results Section describes both the most influential variables and the patterns identified by this analysis.

Data availability

The survey answers analysed in this paper will be made available together with the manuscript.

Code availability

The survey answers analysed in this paper will be made available together with the manuscript.

Received: 5 May 2021; Accepted: 3 August 2021

Published online: 20 September 2021

References

- Zheng, Q. *et al.* Hit-covid, a global database tracking public health interventions to covid-19. *Sci. Data* **7**, 1–8 (2020).
- Desvars-Larrive, A. *et al.* A structured open dataset of government interventions in response to covid-19. *Sci. Data* **7**, 1–9 (2020).
- López, L. & Rodó, X. The end of social confinement and covid-19 re-emergence risk. *Nat. Hum. Behav.* **4**, 746–755 (2020).
- Lai, S. *et al.* Effect of non-pharmaceutical interventions to contain covid-19 in china. *Nature* **585**, 410–413 (2020).
- Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: A model-based analysis. *Lancet. Infect. Dis.* **20**, 669–677 (2020).
- Candido, D. S. *et al.* Evolution and epidemic spread of sars-cov-2 in Brazil. *Science* **369**, 1255–1260 (2020).
- Day, T., Gandon, S., Lion, S. & Otto, S. P. On the evolutionary epidemiology of sars-cov-2. *Curr. Biol.* **30**, R849–R857 (2020).
- Betsch, C. *et al.* Social and behavioral consequences of mask policies during the covid-19 pandemic. *Proc. Natl. Acad. Sci.* **117**, 21851–21853 (2020).
- Gollwitzer, M., Platzer, C., Zwarg, C. & Göritz, A. S. Public acceptance of covid-19 lockdown scenarios. *Int. J. Psychol.* <https://doi.org/10.1002/ijop.12721> (2020).
- Wang, C. *et al.* Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (covid-19) epidemic among the general population in china. *Int. J. Environ. Res. Public Health* **17**, 1729 (2020).
- Liu, W., Yue, X.-G. & Tchounwou, P. B. Response to the covid-19 epidemic: the chinese experience and implications for other countries (2020).
- Abiad, A., Arao, R. M. & Dagli, S. The economic impact of the covid-19 outbreak on developing asia. *Asian Dev. Bank* (2020).
- Fernandes, N. Economic effects of coronavirus outbreak (covid-19) on the world economy. *Available at SSRN* **2**, 3557504 (2020).
- Ivanov, D. Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (covid-19/sars-cov-2) case. *Transp. Res. Part E Log. Transport. Rev.* **136**, 101922 (2020).
- Chen, D.-G., Chen, X. & Chen, J. K. Reconstructing and forecasting the covid-19 epidemic in the united states using a 5-parameter logistic growth model. *Global Health Res. Policy* **5**, 1–7 (2020).
- Clemens, V. *et al.* Potential effects of “social” distancing measures and school lockdown on child and adolescent mental health. *European Child & Adolescent Psychiatry* (2020).
- Jiao, W. Y. *et al.* Behavioral and emotional disorders in children during the covid-19 epidemic. *J. Pediatr.* **221**, 264 (2020).
- Sadati, A. K., B Lankarani, M. H. & Bagheri Lankarani, K. Risk society, global vulnerability and fragile resilience; sociological view on the coronavirus outbreak (2020).
- Mukhtar, S. Psychological health during the coronavirus disease 2019 pandemic outbreak. *Int. J. Soc. Psychiatry* **2**, 0020764020925835 (2020).
- Gao, J. *et al.* Mental health problems and social media exposure during covid-19 outbreak. *PLoS ONE* **15**, e0231924 (2020).
- Hawryluck, L. *et al.* Sars control and psychological effects of quarantine, toronto, canada. *Emerg. Infect. Dis.* **10**, 1206 (2004).
- Wang, Y. *et al.* The impact of quarantine on mental health status among general population in china during the covid-19 pandemic. *Mol. Psychiatry* **2**, 1–10 (2021).
- Fischer, R. *et al.* Rapid review and meta-meta-analysis of self-guided interventions to address anxiety, depression, and stress during covid-19 social distancing. *Front. Psychol.* **11**, 2 (2020).
- Colombo, R. A., Wallace, M. & Taylor, R. An essential service decision model for aba providers during crisis. *Behav. Anal. Pract.* **1**, 2 (2020).
- Seale, H. *et al.* Covid-19 is rapidly changing: Examining public perceptions and behaviors in response to this evolving pandemic. *medRxiv* (2020).
- Yamada, Y. *et al.* Covidistress global survey dataset on psychological and behavioural consequences of the covid-19 outbreak. *Sci. Data* **8**, 1–23 (2021).
- Bailon, C. *et al.* Covidaffect, real-time monitoring of mood variations following the covid-19 outbreak in spain. *Sci. Data* **7**, 1–10 (2020).
- Amat, F., Arenas, A., Falcó-Gimeno, A. & Muñoz, J. Pandemics meet democracy. experimental evidence from the covid-19 crisis in spain. *SocArXiv* (2020).
- Tintori, A. *et al.* Adaptive behavioural coping strategies as reaction to covid-19 social distancing in italy. *Eur. Rev. Med. Pharmacol. Sci.* **24**, 10860–10866 (2020).
- Landi, G., Pakenham, K. I., Boccolini, G., Grandi, S. & Tossani, E. Health anxiety and mental health outcome during covid-19 lockdown in italy: the mediating and moderating roles of psychological flexibility. *Front. Psychol.* **11**, 2195 (2020).
- McBride, O. *et al.* Monitoring the psychological impact of the covid-19 pandemic in the general population: an overview of the context, design and conduct of the covid-19 psychological research consortium (c19prc) study. *PsyArXiv* (2020).
- Freeman, D. *et al.* Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in england. *Psychol. Med.* **2**, 1–13 (2020).
- Chan, H. F., Skali, A., Savage, D. A., Stadelmann, D. & Torgler, B. Risk attitudes and human mobility during the covid-19 pandemic. *Sci. Rep.* **10**, 1–13 (2020).
- Oliver, N., Barber, X., Roomp, K. & Roomp, K. Assessing the impact of the covid-19 pandemic in spain: Large-scale, online, self-reported population survey. *J. Med. Internet Res.* **22**, e21319 (2020) <http://www.jmir.org/2020/9/e21319/>.
- World Health Organization (WHO). Pandemic fatigue. reinvigorating the public to prevent covid-19. Policy framework for supporting pandemic prevention and management (2020). <https://apps.who.int/iris/bitstream/handle/10665/335820/WHO-EURO-2020-1160-40906-55390-eng.pdf>.
- Reicher, S. & Drury, J. Pandemic fatigue? how adherence to covid-19 regulations has been misrepresented and why it matters. *BMJ* **372** (2021). <https://www.bmj.com/content/372/bmj.n137>
- Zerbe, K. J. Pandemic fatigue: Facing the body’s inexorable demands in the time of covid-19. *J. Am. Psychoanal. Assoc.* **68**, 475–478. <https://doi.org/10.1177/0003065120938774> (2020).
- Al-Tammemi, A. B., Tarhini, Z. & Akour, A. A swaying between successive pandemic waves and pandemic fatigue: Where does jordan stand? *Annals of Medicine and Surgery* **65**, 102298 (2021). <https://www.sciencedirect.com/science/article/pii/S20490812100248X>

39. MacIntyre, C. R. et al. Mask use, risk-mitigation behaviours and pandemic fatigue during the covid-19 pandemic in five cities in australia, the uk and usa: A cross-sectional survey. *International Journal of Infectious Diseases* **106**, 199–207 (2021). <https://www.sciencedirect.com/science/article/pii/S1201971221002745>.
40. Seiter, J. S. & Curran, T. Social-distancing fatigue during the covid-19 pandemic: a mediation analysis of cognitive flexibility, fatigue, depression, and adherence to cdc guidelines. *Commun. Res. Rep.* **38**, 68–78. <https://doi.org/10.1080/08824096.2021.1880385> (2021).
41. Ilesanmi, O. S., Bello, A. E. & Afolabi, A. A. Covid-19 pandemic response fatigue in africa: causes, consequences, and counter-measures. *Pan Afr. Med. J.* **37**(Suppl 1), 2 (2020).
42. Petherick, A. et al. A worldwide assessment of covid-19 pandemic-policy fatigue. SSRN (2021). <https://ssrn.com/abstract=3774252>
43. Petersen, E. et al. Covid-19-we urgently need to start developing an exit strategy. *Int. J. Infect. Dis.* **2**, 2 (2020).
44. Kimhi, S., Marciano, H., Eshel, Y. & Adini, B. Resilience and demographic characteristics predicting distress during the covid-19 crisis. *Social Science and Medicine* **265**, 113389 (2020). <https://www.sciencedirect.com/science/article/pii/S0277953620306080>
45. Galasso, V. et al. Gender differences in covid-19 attitudes and behavior: Panel evidence from eight countries. *Proc. Natl. Acad. Sci.* **117**, 27285–27291 (2020).
46. Haischer, M. H. et al. Who is wearing a mask? gender-, age-, and location-related differences during the covid-19 pandemic. *PLOS One* **2**, 3 (2020).
47. Almeida, M., Shrestha, A. & Stojanac, D. The impact of the covid-19 pandemic on women's mental health. *Arch Womens Ment Health* **23**, 741–748 (2020).
48. de Paz, C., Muller, M., Munoz Boudet, A. M. & Gaddis, I. Gender dimensions of the covid-19 pandemic. World Bank Policy Note (2020). <https://openknowledge.worldbank.org/handle/10986/33622>
49. Rodríguez-Rey, R., Garrido-Hernansaiz, H. & Collado, S. Psychological impact and associated factors during the initial stage of the coronavirus (covid-19) pandemic among the general population in spain. *Front. Psychol.* **11**, 1540 (2020). <https://ellisalicante.org/en/covid19impactsurvey>. COVID19ImpactSurvey (2020).
51. Power, K. The covid-19 pandemic has increased the care burden of women and families. *Sustainability: Science, Practice and Policy* **16**, 67–73 (2020).
52. Farre, L., Fawaz, Y., Gonzalez, L. & Graves, J. How the covid-19 lockdown affected gender inequality in paid and unpaid work in spain. IZA Discussion Paper **13434** (2020). <https://ssrn.com/abstract=3643198>.
53. Donato, S. Gender-based violence against women in intimate and couple relationships the case of spain and italy during the covid-19 pandemic lockdown Italian. *Sociol. Rev.* **10**, 869–887 (2020).
54. Cao, W. et al. The psychological impact of the covid-19 epidemic on college students in china. *Psychiatry Res.* **287**, 2 (2020).
55. Benjamin, Y., Tan, N. W., Chew, G. & Lee, K. H. Psychological impact of the covid-19 pandemic on health care workers in Singapore. *Ann. Intern. Med.* **173**, 317–320 (2020).
56. Brooks, S. K. et al. The psychological impact of quarantine and how to reduce it: Rapid review of the evidence. *Lancet* **395**, 912–920 (2020).
57. Morgul, E. et al. Covid-19 pandemic and psychological fatigue in turkey. *Int. J. Soc. Psychiatry* **advance online publication** (2020)
58. Terry, P. C., Parsons-Smith, R. L. & Terry, V. R. Mood responses associated with covid-19 restrictions. *Front. Psychol.* **11**, 3090 (2020).
59. Ceylan, R. F., Ozkan, B. & Mulazimogullari, E. Historical evidence for economic effects of covid-19 (2020). <https://www.rtve.es/noticias/20200612/guia-fases-desescalada/2013020.html>. RTVE (2020).
61. Bruin, J. newestest: command to compute new test. <https://stats.idre.ucla.edu/stata/ado/analysis/> (2011).
62. Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*, 199–213 (Springer, 1998).
63. Hosmer, D. W., Hosmer, T., Le Cessie, S. & Lemeshow, S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat. Med.* **16**, 965–980 (1997).
64. Fayyad, U. & Irani, K. Multi-interval discretization of continuous-valued attributes for classification learning. *Jet Propulsion Laboratory TRS* (1993).
65. Almiñana, M., Escudero, L., Pérez-Martín, A., Rabasa, A. & Santamaría, L. A classification rule reduction algorithm based on significance domains. *TOP* **22**, 397–418 (2014).
66. Quinlan, J. *Introduction of decision trees machine learning* Vol. 1, 86–106 (Kluwer Academic Publishers, Boston (NL), 1986).

Acknowledgements

This work is partially supported by grants FONDOS SUPERA COVID-19 Santander-CRUE, 2020–2021, Fundación BBVA a equipos de investigación científica SARS-CoV-2 COVID-19, IA4COVID19 2020–2022 and the Valencian Government.

Author contributions

M.M.-G. Contributed to the study design, performed the data processing and logistic regression modeling. A.R. Contributed to the study design, performed the modeling and automatic pattern extraction. X.B. Contributed to the study design, performed the data processing and logistic regression modeling. K.P. Contributed to the study design, performed the modeling and automatic pattern extraction. K.R. Contributed to the design and implementation of the online survey and raw data processing. N.O. Contributed to the design and implementation of the online survey, the study design, the definition of the research questions, the coordination of all the tasks, data analysis and insights extraction. All authors contributed to the drafting of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-97645-1>.

Correspondence and requests for materials should be addressed to N.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021