I’m alone but not lonely. U-shaped pattern of perceived loneliness during the COVID-19 pandemic in the UK and Greece

Abstract

Objectives: In the past months, many countries have adopted varying degrees of lockdown restrictions to control the spread of the COVID-19 virus. According to the existing literature, some consequences of lockdown restrictions on people’s lives are beginning to emerge. To inform policies for the current and/or future pandemics, particularly those involving lockdown restrictions, this study adopted a data-driven Machine Learning approach to uncover the short-term effects of lockdown on people’s physical and mental health. Study design: An online questionnaire about participants’ health and life during the pandemic was launched on 17 April 2020 and was completed by 2,276 people from 66 different countries. Methods: Focusing on the UK sample (N=382), 10 aggregated variables representing the participant’s living environment, physical and mental health were used to train a RandomForest model to predict the week of survey completion. Results: Using an index of importance to identify the best predictor among the 10 variables, self-perceived loneliness was identified as the most influential variable. Subsequent statistical analysis showed a significant U-shaped curve for loneliness levels, with a decrease during the 4th and 5th lockdown weeks. The same pattern was replicated in the Greek sample (N = 129). Conclusions:
This suggests that for the very first period of time, the adopted lockdown measures affected people’s evaluation of their social support, leading to a decreased sense of loneliness.

*Keywords:* machine learning; COVID-19; lockdown; loneliness; global study; mental health
1. Introduction

The 2019 SARS-CoV-2 (COVID-19) outbreak was declared as a pandemic by the World Health Organisation (WHO) on 11 March 2020. The number of positive cases worldwide at the time was 179,111 and deaths, 7,426 [1].

Globally, the months that followed saw a surge in the number of deaths and infection rates, which put further strain on the sanitary and economical balance of several countries. Fast forward to September 13th 2020, the total number of confirmed COVID-19 cases at 28,637,952 and 917,417 deaths have since been recorded [2].

Expert concern for the mental health consequences of the current pandemic stems from the evidence that was obtained during smaller epidemics, such as SARS (severe acute respiratory syndrome), MERS (Middle East respiratory syndrome-related coronavirus), H1N1, and Ebola. From these previous health emergencies, short- and long-term effects on the healthcare workers’ mental health, such as post-traumatic stress disorder (PTSD) [3, 4], depression [5, 6], anxiety [5], stress and burnout [3] symptoms were common [7]. There is evidence that healthcare workers are distressed from the epidemics, during and after emergencies, and that these effects also extend to the general population in the form of severe anxiety, post-traumatic stress disorder, depression and increased rates of substance abuse [8, 9, 10, 11, 12, 13]. Although some promising results from vaccine trials are starting to emerge, the novel and highly infective virus continues to force governments around the world to limit people’s movements and, in some cases, re-adapt lockdown restrictions once again, as in the case of the UK on September 22nd, 2020.

Closing schools and universities, shutting non-essential businesses, enfor-
ing working from home policies and online teaching, struggling with financial
difficulties and leaving the house only for necessities have fuelled genuine and
perceived health threats that have rapidly become ubiquitous for large popu-
lations worldwide. While restrictions have helped flatten the infection curve,
legitimate concerns about the physical and mental health consequences have
been raised. As such, this pandemic, as an extreme global stressor, has pro-
vided an unprecedented opportunity for researchers to investigate how several
aspects of our personal life, and specifically our mental health, are affected
by prolonged isolation and restrictions. Social isolation is one known threat
to mental and physical well-being [14, 15] and an established risk factor for
mortality [16, 17, 18]. Social isolation is associated with poor sleep quality
[19] and with an increased risk of cognitive decline [20]. The fact that our
perception of self is ingrained in the social comparison with others [21] sug-
gests that social isolation may not be an ideal situation for the development
of our identity either. Latest COVID-19 studies of the first weeks of lockdown
have already documented psychological distress, such as depression, anxiety,
post-traumatic stress and insomnia in Italy [22, 23, 24] and China [25, 26],
the two countries most severely hit by COVID-19 at the beginning of the
pandemic, as well as Austria [27] and Switzerland [28]. In fact, with some
preliminary results on COVID-19 restrictions, this paper aims to add a piece
of knowledge to the existing literature to provide a scientific contribution and
help governments in the design of future possible lockdowns. Against this
backdrop of existing psychological consequences from lockdown, this study
focuses on the physical and psychological constructs that best predict the
time spent in lockdown (TIL).
2. Methods

2.1. Questionnaire

A 20-minute online survey (available in 7 languages) was administered through the website link blinded for review between 17 April 2020 and 10 July 2020 to participants aged 18 years and above who had access to the survey link. This was distributed using various social media channels (email, LinkedIn, Whatsapp, Instagram, Facebook and Reddit). The survey was designed by group blinded for review in order to explore participants’ moods and behaviours. The battery of questionnaires consisted of 359 questions assessing 13 main domains: social suspicions, schizotypal traits, physical health, sleep quality, aggression, empathy, anxiety, depression, worries and stress, loneliness, parenting style, Special Educational Needs and demographic information (see pre-registration link blinded for review for more details). The study was approved by the XXX IRB.

2.2. Participants

Participants for the study were recruited through convenience sampling and, eventually, a total of 2,276 people (aged 18 and above) from 66 countries completed the survey during lockdown. Respondents who did not give consent to treat their data (N = 32), with incomplete (N = 712) or missing data (N = 294), or who could not complete the survey within two days from their enrollment (N = 76) were excluded. To train the RandomForest, we chose not to consider the participants who took more than one day because the process required the model to find patterns of dependency between the features and the amount of time spent in lockdown. Considering the fact that,
in our hypothesis, the time in lockdown played a role in determining the variability of the selected features, by considering only the participants who completed the survey within the same day, we aimed at reducing possible confounds. Furthermore, participants who completed the survey from a country that was different from the one they were a resident of were excluded from the study (N = 132). Considering the variety of lockdown measures across the world, this criterion was adopted in order to reduce possible confounds given by the type of restrictions adopted by individual countries. Another possible confound came from the fact that not all the governments decided to adopt lockdown restrictions against the pandemic and, when they did, different countries entered the lockdown on different dates. For these reasons, among the countries that adopted these restrictions, the new variable "Weeks in lockdown" - the time of survey completion - was computed for each participant. Thus, participants were grouped and compared regardless of the specific date in which their countries decided to adopt restrictions, but uniquely by the amount of time spent in lockdown. Within this pool of data, the UK and Greece samples were selected for the analysis conducted in this study. Three main reasons drove the choice of using data from these two countries: a) the sample sizes (≥100 cases); b) the existence of a clear date of lockdown beginning; and c) the same time span (weeks in lockdown) covered. To maintain the coverage on the same time period, UK participants that completed the survey after week 9 of lockdown were excluded from the study (N = 40). To summarise, the UK sample consisted of 382 participants (Gender: Female = 302, Male = 71, Non-binary = 4, Prefer not to say = 2, Self-identified = 3; Age: mean = 37.18; SD = 13.15), while the Greek sample
counted 129 participants (Gender: Female = 92, Male = 37; Age: Mean = 36.08, SD = 10.79) (see Table 1).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Week 6</th>
<th>Week 7</th>
<th>Week 8+</th>
<th>TOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>38</td>
<td>78</td>
<td>75</td>
<td>69</td>
<td>97</td>
<td>25</td>
<td>382</td>
</tr>
<tr>
<td>Greece</td>
<td>7</td>
<td>86</td>
<td>18</td>
<td>17</td>
<td>1</td>
<td>0</td>
<td>129</td>
</tr>
</tbody>
</table>

Table 1: Distribution of participants from the UK and Greece by week.

2.3. Data Analysis

From the dataset, 10 variables capturing participants’ living environment, mental and physical health were selected (see Table 2). This procedure aimed to remove the noise in the dataset by discarding the non-informative variables. All the variables for which scores could not show a variation caused by the amount of time in lockdown were excluded. Some examples of these variables are gender, ethnicity, dimensions of one’s house, and other similar variables. Moreover, considering the criteria for which data from participants could potentially have at least one NA data were not considered for the analysis, the next step of the selection was to optimise the number of participants by not considering the variables that had a large amount of NA data.

In order to investigate the role of time spent in lockdown on modulating the effects of lockdown restrictions (study pre-registration: pre-registration link blinded for review), the study consisted of two parts. In the first, we aimed at understanding which aspect was most sensitive to time. In other words, which part of our physical and mental life was affected the most by the time in lockdown. In the second, the interest was on how this most sensitive aspect was modulated by the time.
<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild activity difference</td>
<td>Difference between days of mild physical activity post- and pre- COVID-19 lockdown.</td>
<td><em>International Physical Activity Questionnaire – Short Form (IPAQ-SF, 6-items)</em> [29]</td>
</tr>
<tr>
<td>Mild activity time difference</td>
<td>Difference between minutes of mild physical activity post- and pre- COVID-19 lockdown.</td>
<td><em>International Physical Activity Questionnaire – Short Form (IPAQ-SF, 6-items)</em> [29]</td>
</tr>
<tr>
<td>Moderate activity difference</td>
<td>Difference between days of moderate physical activity post- and pre- COVID-19 lockdown.</td>
<td><em>International Physical Activity Questionnaire – Short Form (IPAQ-SF, 6-items)</em> [29]</td>
</tr>
<tr>
<td>Sleep quality</td>
<td>Self-reported sleep quality and quantity, where higher scores reflect better sleep quality.</td>
<td><em>Pittsburgh Sleep Quality Index</em> (2-items) [30], <em>Subjective and Objective Sleepiness Scale</em> [32]</td>
</tr>
<tr>
<td>Empathy</td>
<td>Self-reported affective, cognitive, and somatic empathy, where higher scores reflect higher empathy.</td>
<td><em>Cognitive, Affective, Somatic Empathy Scale (CASES, 30-items)</em> [33]</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Higher scores reflect higher anxiety.</td>
<td><em>General Anxiety Disorder-7 (GAD-7)</em> [34]</td>
</tr>
<tr>
<td>Depression</td>
<td>Higher scores reflect higher depression.</td>
<td><em>Patient Health Questionnaire-9 (PHQ-9, 9-items)</em> [35]</td>
</tr>
<tr>
<td>Perceived loneliness</td>
<td>Higher scores reflect higher perceived loneliness.</td>
<td><em>Loneliness Questionnaire (LQ, 20-items)</em> [36]</td>
</tr>
<tr>
<td>Living conditions/environment</td>
<td>Higher scores reflect more chaotic home environments.</td>
<td><em>Chaos, Hubbub, and Order Scale and Health Risk Behaviors (CHAOS, 6-items)</em> [37]</td>
</tr>
<tr>
<td>Beliefs</td>
<td>Perceived effectiveness of government guidelines on social distancing, schools closing, face masks and gloves as protection. Higher scores reflect stronger beliefs.</td>
<td>Summed 9-items on COVID-19 beliefs</td>
</tr>
</tbody>
</table>

Table 2: Scores that are computed to quantify participants’ mental and physical health and living environment during lockdown.

Identification of the most influential variable

Without any available literature to guide our hypothesis in identifying the variable that is most influenced by TIL, we adopted a data-driven Machine Learning approach where a RandomForest [38] regression model was trained to predict the week in which each participant completed the survey, based on the total scores of the 10 selected variables. The model creates an ensemble of decision trees based on the predictive information of the input variables.
The performance of the model was evaluated based on Mean Squared Error (MSE). The data used to train the RandomForest model were those of the 382 UK residents who were in the UK at the time of participation in this study. Initially, the dataset was partitioned into train (75% of participants) and test (25% of participants). The training process was repeated and evaluated several times on different randomized folds of the train dataset to optimize the number of decision trees and rank the variables based on their importance. A Borda count [39] was then computed on the rankings of variables obtained from each training iteration to identify the most important variable to predict the week of survey completion. The optimal number of decision trees that emerged from the training was 50. The final model, with the optimal number of trees, was then trained on the whole train partition and evaluated on the test partition. The adopted training scheme is standardized and was derived from bioinformatics applications that are used to identify clinical biomarkers from genetic data [40].

Statistical validation

In the second part of the study, we used a Kruskal-Wallis test to assess whether the most important variable (identified by the RandomForest model) significantly changes during the lockdown from weeks 3 to 7. In case of significant results, we adopted post-hoc Kruskal-Wallis tests to compare pairwise the 3rd week with the 4th to 7th weeks. The Bonferroni method was used to correct the significance level for multiple comparisons. In conducting statistical analyses, we first focused on the same set of participants used to train the Random Forest model, then we validated results on the dataset of participants from Greece.
3. Results

MSE on the training and the test partitions was 1.33 and 1.94 respectively and the feature with the highest importance was perceived loneliness (see Figure 1).

![Figure 1: Average importance of the selected variables.](image)

Notably, scores of perceived loneliness decreased during weeks 3 to 5 after lockdown and subsequently increased in the following weeks, returning to the initial values (see Figure 2). The Kruskal-Wallis test on the data of UK participants from the 3rd to 7th week confirmed that at least one week was statistically different from the others ($H=12.86, p=0.012$). We then compared the 4th, 5th, 6th and 7th week with the 3rd. Significant differences were found for the 4th ($H=11.360, p=0.001$) and 5th ($H=7.077, p=0.008$) week,
but not for the 6th (H=4.011, p=0.045) and the 7th (H=0.368, p=0.544) week. The same procedure was repeated on participants from Greece, focusing only on the 3rd to 6th weeks, as only one participant completed the survey during the 7th week. The results confirmed that perceived loneliness significantly changes over weeks 3 to 6 (H=8.27, p=0.041), with a significant difference between the 3rd and the 5th weeks (H=7.6, p=0.006). The difference between the 3rd and the 4th weeks (H = 3.87, p=0.049) failed to survive the Bonferroni correction. No difference was found between the 3rd and the 6th week (H=0.68, p=0.408).

Figure 2: Distribution of perceived loneliness scores for each week for participants from the UK (left) and Greece (right).
4. Discussion

The aim of this study was to assess the impact of lockdown restrictions on people’s mental and physical health. Although we adopted a rigorous methodology to train the predictive model, the achieved performances on train and test partitions are low: this outcome reflects the complexity of the psycho-social mechanisms that were at play during the lockdown period. While the questionnaires aimed at quantifying a broad range of the aspects of interest, other aspects may not have been observed. Additionally, this study investigated the temporal variations of these mechanisms, whose effects on the observed variables might be even more difficult to identify. That said, one advantage of the machine learning approach is that it permits the identification of variables that are more sensitive to the time spent in lockdown, rather than a focus on predictive capability.

The low performance does not affect the reliability of the ranking of the variables, which identified perceived loneliness as the most sensitive variable. Perceived loneliness in the UK decreased during the first 4-5 weeks after the start of the lockdown, before returning to initial values afterwards. The pattern was replicated in the Greek sample, albeit in a smaller group of participants. This confirms that perceived loneliness does capture a sensitive decrease during weeks 4-5 since the start of lockdown.

These results are somewhat surprising. In the emerging literature about COVID-19, the number of friends and one’s social support seem to play a protective role against the effect of lockdown on loneliness [41, 42]. Counter-intuitively, from our study it emerged that, even though a large part of the global population was not able to see their close friends, partner and family,
levels of perceived loneliness interestingly decreased during the initial period of lockdown. The dissociation between the objective and the subjective degrees of social support has been largely discussed in the existing literature [43, 44]. It is believed that perceived social isolation represents a quantitative or (more often) qualitative mismatch between an individual’s need for social support and the subjective evaluation of the social support that is obtained [45]. In other words, the feeling of loneliness, resulting from the perception of social isolation, seems to depend, more than on an objective condition of isolation, on a cognitive evaluation and perception of the social environment.

In the existing literature, the feeling of loneliness emerged to be connected to the concept of Self [46], the person’s cognitive functioning [44] and, in general, the mental and physical well-being of the individual [17, 16, 15, 47]. For instance, lonely people are more likely to suffer from depression [48, 49], Alzheimer’s disease [50, 49], alcoholism [51, 52, 49], suicide [53, 49], personality disorders [49] and sleep problems [49]. It is not yet clear the reason behind the results that emerged from this study, but some hypotheses can be advanced. For instance, considering the definition of loneliness as a mismatch between desired and obtained social support, the decrease of its levels in the first weeks of lockdown could signify that people in that period of time were receiving the desired social support or even more of it in terms of quantity, or that it is higher in quality. On the other hand, loneliness could decrease as a result of a drop in the standards used for evaluating the received social support. In times of danger, this could be an adaptive feature, for it could facilitate behaviours of affiliation among people of the same group. As a matter of fact, facing an external threat has the short-term
effect of increasing cohesiveness among the members of a group [54, 55]. The threat of external dangers (such as invaders) has often been used from past and present political leaders in order to rule their countries and to increase the sense of community among the population. It is possible that not only personified external threats like foreigners trying to invade one’s country, but also environmental dangers, such as the COVID-19 pandemic, can directly modulate the degree of cohesiveness among the members of a group, in this case an entire population of a country. As a matter of fact, this increased cohesion could have had a role in the initial decrease in levels of perceived loneliness that we observed among people from the UK and Greece. Even though no certain explanation can be given for understanding the observed patterns of perceived loneliness, the findings of this study support the idea that social isolation (as the objectively low social support) and loneliness (as the subjectively low social support) are different concepts, not necessarily linked to each other, as philosophers in the past centuries have largely pointed out. Having observed that lockdown restrictions have short-term effects on people’s feeling of loneliness and not knowing the real meaning behind these observed patterns, in our opinion, the design of possible future lockdown measures should be accompanied by the consideration of the role played by real and perceived social support for people’s physical and mental well-being.
Author contribution

BLINDED FOR REVIEW All authors read and agreed to the published version of the manuscript.

Conflicts of interest

The authors declare no conflict of interest.

Ethics

Ethical approval for the COVID-19 Social Study was granted by the BLINDED FOR REVIEW. The study is GDPR compliant.

Funding

This research is supported by BLINDED FOR REVIEW.

References


